AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LVIII
1936

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

INDEX.

A TOPOLOGICAL PROOF OF THE RIEMANN-ROCH THEOREM ON AN ALGEBRAIC CURVE.

By Oscar Zariski.

The purpose of this paper is to develop a *topological theory of linear series on an algebraic curve.* Our chief tool is the *symmetric topological product* of the Riemann surface of the curve: its use was clearly indicated by the involutorial character of linear series and plays here the same rôle as the notion of the direct topological product plays in Lefschetz's topological theory of algebraic correspondences (see S. Lefschetz (3); also Todd (5), Zariski (6), chapter VI and Appendix B). It will be seen from the following exposition that the use of symmetric products in the geometry on an algebraic curve leads very rapidly to the very heart of the theory, including the central theorem of Riemann-Roch.

1. By the *symmetric n-th product* of a complex K is meant the topological space of the *unordered* sets of n points of K. This space, which we shall denote by K^n, is in $(1, n!)$ correspondence with the direct topological product of n complexes homeomorphic to K. The involution of sets of $n!$ points defined by this correspondence on the direct product is generated by a group G of automorphisms T_ν, isomorphic with the symmetric group of substitutions on n letters. From this point of view symmetric products were studied recently by M. Richardson,[4] who gave a convenient simplicial subdivision of K^n, outlined a general procedure for the determination of its Betti numbers and determined explicitly these numbers in the cases $n = 2$ and $n = 3$.

In this paper we shall be concerned with the symmetric n-th product R^n of a Riemann surface R of genus p. If $p = 0$, R is a 2-sphere and R^n is the space of all unordered sets of n (distinct or coincident) complex numbers $z_1, z_2, \cdots, z_n$, including $z = \infty$. With any such set we associate the coefficients, determined to within a factor of proportionality, of the polynomial $a_0 z^n + a_1 z^{n-1} + \cdots + a_n$ of which the numbers of the set are the roots. If k of the numbers of the set coincide with $z = \infty$, then $a_0 = a_1 = \cdots = a_{k-1} = 0$. It follows that *in the case $p = 0$ R^n is homeomorphic with the complex projective n-dimensional space.* Since the manifold condition is of a local character, it follows that *R^n is an absolute manifold* also if $p > 0$.

We make a few preliminary remarks concerning the boundary and orientation of the symmetric product of a 1-cell and of a 2-cell.

Let the 1-cell E_1 be the segment $(0 - 1)$ of the real line x. It is seen immediately that $E_1{}^n$ is homeomorphic with the following point set in the Euclidean space $(x_1, x_2, \cdots, x_n)$: $0 < x_1 \leqq x_2 \leqq \cdots \leqq x_n < 1$. This point set is a simplex σ_n whose $n+1$ faces lie on the hyperplanes $x_1 = 0$, $x_n = 1$, $x_i = x_{i+1}$, $(i = 1, 2, \cdots, n-1)$. Of these faces only the first two arise from the boundary of E_1, while the remaining faces arise from the n-tuples of points of E_1 in which two or more points coincide. In this circumstance lies the reason of the fact that the n-th symmetric product of a 1-sphere is only a relative manifold (a manifold with a boundary). Thus, for $n = 2$, it is the Moebius ring.

On the contrary, the n-th symmetric product $E_2{}^n$ of a 2-cell E_2 is a $2n$-cell whose boundary arises exclusively from the boundary of E_2, its points corresponding to those n-tuples of points of the closed E_2 which have at least one point on the boundary of E_2. To see this it is sufficient to introduce in E_2 a complex parameter z and to apply the method of symmetric functions of the roots $z_1, z_2, \cdots, z_n$ used above. For this reason the symmetric product of an absolute manifold M_2 is an absolute manifold.

For a given orientation of E_2 a corresponding orientation of $E_2{}^n$ may be defined as follows. Let $(P_1, P_2, \cdots, P_n)$ be a set of n distinct points of E_2 and let (ξ_i, η_i) be an indicatrix of the oriented E_2 at P_i. The symbol $(\xi_1, \eta_1; \xi_2, \eta_2; \cdots; \xi_n, \eta_n)$ defines in an obvious manner an indicatrix at the point $P \equiv (P_1, P_2, \cdots, P_n)$ on $E_2{}^n$ and hence an orientation of $E_2{}^n$. *This orientation is independent of the order of the points P_i of the n-tuple,* because any permutation of the points P_i induces an even permutation of the elements of the indicatrix at P. As a corollary it follows that *R^n is an orientable manifold.*

An analogous consideration for the orientation of $E_1{}^n$ leads to an indicatrix $(\xi_1, \xi_2, \cdots, \xi_n)$ which is altered by an odd permutation of the points P_i. Since on a 1-sphere H_1 it is possible to deform into each other any two *ordered* n-tuples of distinct points of H_1 without crossing n-tuples having coincident points, it follows that the relative manifold $H_1{}^n$ is non orientable.

2. *Cycles and minimal bases on R^n.* It has a sense to speak of the symmetric product also when the factors $K_1, K_2, \cdots$ are distinct subcomplexes of a given complex K or chains on K. When these subcomplexes or chains have cells in common their symmetric product is in general different from the direct topological product. While keeping the usual notation $K_1 \times K_2 \times \cdots$ for the direct topological product, we shall denote the symmetric product by $K_1K_2 \cdots$.

Let $M_{2n} = R \times R \times \cdots \times R$ be the direct topological product of n

factors R and let us consider the correspondence $(1, n!)$ between R^n and M_{2n}. We denote by $\phi(P')$ the homologue P on R^n of a point P' on M_{2n} and by $\psi(P)$ the set of $n!$ points on M_{2n} which correspond to a given point P on R^n. The operations ϕ and ψ transform chains into chains and preserve boundary relations. In particular, they transform cycles into cycles and homologous cycles into homologous cycles (see M. Richardson (4)). If C is a chain on R^n, then $\phi\psi(C) = n!C$.

Let $\Gamma_k^1, \Gamma_k^2, \cdots$ be a maximal set of independent k-cycles on M_{2n} and let $\Delta_k^i = \phi(\Gamma_k^i)$ be the corresponding cycles on R^n. Let Δ_k be an arbitrary k-cycle on R^n and let $\Gamma_k = \psi(\Delta_k)$ be the corresponding cycle on M_{2n}. We have $t\Gamma_k \sim t_1\Gamma_k^1 + t_2\Gamma_k^2 + \cdots$. Operating by ϕ and observing that $\phi(\Gamma_k) = n!\Delta_k$, we obtain: $n!t\Delta_k \sim t_1\Delta_k^1 + t_2\Delta_k^2 + \cdots$. Hence *every k-cycle on R^n depends on the cycles Δ_k^i.*[1]

Let $\delta_1, \delta_2, \cdots, \delta_{2p}$ be a minimal base of 1-cycles for homology $\sim$ on R. It is well known that the following k-cycles on M_{2n} form a minimal base for homology $\sim$:

$$\Gamma_k^{(i)} = \Gamma_k^{i_1, i_2, \cdots, i_\beta} = M_{2\alpha} \times \delta_{i_1} \times \delta_{i_2} \times \cdots \times \delta_{i_\beta} \times P_1 \times P_2 \times \cdots \times P_\gamma,$$
$$2\alpha + \beta = k, \qquad \alpha + \beta + \gamma = n,$$

where $M_{2\alpha}$ is the direct product of α factors R and $P_1, P_2, \cdots, P_\gamma$ are arbitrary fixed points on R. Of what nature are the corresponding cycles $\phi(\Gamma_k^{(i)})$ on R^n? Let us first consider the case in which the indices $i_1, i_2, \cdots, i_\beta$ are all distinct. In this case the correspondence between $\phi(\Gamma_k^{(i)})$ and $\Gamma_k^{(i)}$ is $(1, \alpha!)$ and the cycle $\Gamma_k^{(i)}$ is transformed into itself by the $\alpha!$ automorphisms T_ν of M_{2n} (see section 1) which correspond to the permutations of the α factors R in $M_{2\alpha}$. *These automorphisms preserve the orientation of $\Gamma_k^{(i)}$*, since the permuted factors are of even dimension. It follows that in $\phi(\Gamma_k^{(i)})$ each oriented cell is repeated $\alpha!$ times and that consequently,

$$\phi(\Gamma_k^{i_1, i_2, \cdots, i_\beta}) = \alpha!\,\Delta_k^{i_1, i_2, \cdots, i_\beta} = \alpha!\,\Delta_k^{(i)},$$

where $\Delta_k^{(i)}$ is a cycle on R^n which as a locus of points consists of all the unordered n-tuples of points of R having one point on δ_{i_1}, one point on $\delta_{i_2}, \cdots,$ one point on δ_{i_β} and γ points coincident with $P_1, P_2, \cdots, P_\gamma$. In our notation for symmetric products we have

(1) $$\Delta_k^{(i)} = \Delta_k^{i_1, i_2, \cdots, i_\beta} = R^\alpha \delta_{i_1}\delta_{i_2} \cdots \delta_{i_\beta} P_1 P_2 \cdots P_\gamma,$$
where
(2) $$2\alpha + \beta = k, \qquad \alpha + \beta + \gamma = n.$$

---

[1] In M. Richardson (4) the above proof is developed along strictly combinatorial lines.

The orientation of $\Delta_k^{(i)}$ is determined by the orientation of $\Gamma_k^{(i)}$ and depends on the orientation of R and on the orientation and on the order of the cycles $\delta_{i_1}, \delta_{i_2}, \cdots, \delta_{i_\beta}$.

Let now $\Gamma_k^{(i)}$ be a cycle with repeated indices, for instance, let $i_1 = i_2$. The involutorial automorphism T of M_{2n} which interchanges the factors δ_{i_1} and δ_{i_2} transforms $\Gamma_k^{(i)}$ into $-\Gamma_k^{(i)}$. Hence assuming a simplicial subdivision s of M_{2n} such that $\phi(s)$ is a simplicial subdivision of R^n (see M. Richardson (4)), we will have for any simplex σ_k of $\Gamma_k^{(i)}$, $\phi(\sigma_k) = -\phi(T\sigma_k)$. It follows that $\phi(\Gamma_k^{(i)})$ *vanishes identically as a chain on* R^n.

We have thus proved that *any k-cycle on R^n depends on the cycles $\Delta_k^{i_1, i_2, \cdots, i_\beta}$ given by* (1), *the indices $i_1, i_2, \cdots, i_\beta$ being all distinct.* Here the order of the indices affects only the sign of the cycle, and therefore the distinct cycles correspond to the unordered sets of indices.

From now on we shall assume that the cycles δ_i form a canonical set of retrosections of R and that their orientations are such that the following intersection formulas hold:

$$(3) \qquad (\delta_i \cdot \delta_{i+p}) = -(\delta_{i+p} \cdot \delta_i) = +1, \qquad (i = 1, 2, \cdots, p);$$

$$(3') \qquad (\delta_i \cdot \delta_j) = 0, \text{ if } |i - j| \neq p.$$

The intersection numbers of the cycles $\Delta_k^{(i)}$ and $\Delta_{2n-k}^{(j)}$ of complementary dimensions on R^n evidently coincide with those of the cycles $\Gamma_k^{(i)}$ and $\Gamma_{2n-k}^{(j)}$ on the direct product M_{2n}. Using the well-known formulas giving the intersection numbers of cycles on direct products (see Lefschetz (2), p. 243), we find in the present case the following intersection formulas:

$$(4) \qquad (\Delta_k^{i_1, i_2, \cdots, i_\beta} \cdot \Delta_{2n-k}^{j_1, j_2, \cdots, j_{\bar\beta}}) = 0, \text{ if } \beta \neq \bar\beta;$$

$$(5) \quad (\Delta_k^{i_1, i_2, \cdots, i_\beta} \cdot \Delta_{2n-k}^{j_1, j_2, \cdots, j_\beta}) = (-1)^{[\beta(\beta-1)/2]} (\delta_{i_1} \cdot \delta_{j_1}) (\delta_{i_2} \cdot \delta_{j_2}) \cdots (\delta_{i_\beta} \cdot \delta_{j_\beta})$$

From the fact that the intersection numbers are different from zero only for pairs of associated cycles $\Delta_k^{i_1, i_2, \cdots, i_\beta}$ and $\Delta_{2n-k}^{i_1+p, i_2+p, \cdots, i_\beta+p \pmod{2p}}$, it follows that *the cycles $\Delta_k^{(i)}$ are independent.* Moreover *these cycles form a minimal base for homology $\approx$,* since the above intersection numbers are all 0 or ± 1.

To calculate the Betti number R_k of R^n we observe that the number β of indices in any k-cycle $\Delta_k^{i_1, i_2, \cdots, i_\beta}$ must satisfy the inequalities

$$(6) \qquad \beta \leqq k, \qquad \beta \leqq 2n - k, \qquad \beta \leqq 2p.$$

The first inequality follows from the relation $2\alpha + \beta = k$. The second is

the dual of the first and follows directly by combining the relations $2\alpha + \beta = k$ and $\alpha + \beta + \gamma = n$, getting $2\gamma + \beta = 2n - k$. The third inequality follows from the fact that the indices $i_1, i_2, \cdots, i_\beta$ are distinct integers $\leqq 2p$. We also observe that the difference $k - \beta$ is even. Hence if ρ denotes the smallest of the three integers k, $2n - k$, $2p$, then

$$\left.\begin{aligned} R_k &= 1 + \binom{2p}{2} + \cdots + \binom{2p}{\rho}, \text{ if } k \text{ is even} \\ R_k &= \binom{2p}{1} + \binom{2p}{3} + \cdots + \binom{2p}{\rho}, \text{ if } k \text{ is odd} \end{aligned}\right\} \rho < 2p,$$

and

$$\left.\begin{aligned} R_k &= 1 + \binom{2p}{2} + \cdots + \binom{2p}{2p}, \text{ if } k \text{ is even} \\ R_k &= \binom{2p}{1} + \binom{2p}{3} + \cdots + \binom{2p}{2p-1}, \text{ if } k \text{ is odd} \end{aligned}\right\} \rho = 2p,$$

i. e., if $\rho = 2p$, then $R_k = 2^{2p-1}$ whether k is even or odd.

3. *Linear series $g_n{}^r$ as cycles on R^n. Theorem of Clifford.*[2] *Let R be* the Riemann surface of an algebraic curve f, $f(x, y) = 0$, of genus p. We consider on f a linear series $g_n{}^r$, i. e., a series ∞^r of sets of n points of the curve which are either the sets of level of a linear system of rational functions on f,

$$t_0 + t_1 g_1(x, y) + \cdots + t_r g_r(x, y),$$

or differ from these sets by any number of fixed points. A set of the $g_n{}^r$ is represented by a point of the symmetric product R^n and the $g_n{}^r$ is given on R^n by a certain $2r$-cycle Γ_{2r}. We express this cycle in terms of the cycles $\Delta_{2r}^{(i)}$ of the base:

$$\Gamma_{2r} \sim \epsilon\, \Delta_{2r} + \Sigma\, \epsilon_{i_1, i_2} \Delta_{2r}^{i_1, i_2} + \cdots + \Sigma\, \epsilon_{i_1, i_2, \ldots, i_{2\rho}} \Delta_{2r}^{i_1, i_2, \ldots, i_{2\rho}}, \tag{7}$$

where $\Delta_{2r} = R^r P_1 P_2 \cdots P_{n-r}$ and where ρ is the smallest of the three numbers p, r, $n - r$ (see the formulas (6), where β and k should be replaced by 2ρ and $2r$ respectively). It should be understood that the above sum contains one and only one term for each unordered set of indices. We agree that the coefficients $\epsilon_{(i)}$ change sign for odd permutations of the indices and are unaltered by even permutations. Then the order of the indices in any cycle $\Delta_{2r}^{(i)}$ is immaterial. For the determination of the coefficients $\epsilon_{(i)}$ we use two properties of a $g_n{}^r$ which follow directly from the definition: a) There exists one and only one set in the $g_n{}^r$ which contains r generic preassigned points of the curve f (*involutorial property*); b) the sets of the $g_n{}^r$ are in

[2] *Notice to the reader:* All homologies in this and in the following section are weak homologies (with allowed division). ***For printing purposes we shall use from now on, without fear of confusion, the symbol $\sim$ instead of $\approx$.***

$(1, 1)$ correspondence with the ratios of the parameters t_i, i. e., with the points of a linear complex space S_r (*rationality of the* $g_n{}^r$). From the involutorial property it follows that if $Q_1, Q_2, \cdots, Q_r$ are generic points of f, then Γ_{2r} intersects in one point only the cycle $\Delta_{2(n-r)} = R^{n-r}Q_1Q_2 \cdots Q_r$. Let $Q \equiv (Q_1, Q_2, \cdots, Q_n)$ be the common point of the two cycles and let u_i be a uniformizing parameter on R in the neighborhood of Q_i. Then $u_1, u_2, \cdots, u_n$ are local coördinates of R^n at the point Q and the equations of the loci Γ_{2r} and $\Delta_{2(n-r)}$ in the neighborhood of Q will be respectively:[3]

$$u_{r+j} = f_j(u_1, u_2, \cdots, u_r),$$

$$[j = 1, 2, \cdots, n - r, f_j\text{—a regular function at } (Q_1, Q_2, \cdots, Q_r)];$$

and

$$u_i = \text{const.} \qquad (i = 1, 2, \cdots, r).$$

From these equations it is seen that we are dealing with two analytical varieties whose tangent spaces are independent, and therefore the intersection number of the two cycles is ± 1. Assuming for R and for Γ_{2r} the intrinsic orientation determined by the analytical character of these varieties (Lefschetz (3), Zariski (6), p. 102), we get the sign $+$. Hence $(\Gamma_{2r} \cdot \Delta_{2(n-r)}) = +1$ and since, by (4) and (5), $(\Gamma_{2r} \cdot \Delta_{2(n-r)}) = \epsilon$, it follows $\epsilon = +1$.

We now use the property b). Since the Betti numbers of odd dimensions of S_r all vanish and since Γ_{2r}, as a locus of points, and S_r are homeomorphic, it follows that *the intersection of* Γ_{2r} *with any cycle of odd dimension on* R^n *is* ~ 0 *on* Γ_{2r} *and hence a fortiriori* ~ 0 *on* R^n. *In particular we have the following homologies:*

$$(7') \qquad \Gamma_{2r} \cdot \Delta^j{}_{2n-1} \sim 0, \qquad (j = 1, 2, \cdots, 2p).$$

To make use of these homologies we have only to express the intersection cycles $\Delta^{(i)}_{2r} \cdot \Delta^j{}_{2n-1}$ in terms of the basic cycles $\Delta^{(i)}_{2r-1}$.

We consider the intersection of $\Delta^j{}_{2n-1}$ with a given cycle $\Delta^{i_1, i_2, \cdots, i_{2m}}_{2r}$ and we examine separately two cases, according as j does not or does coincide with one of the numbers $i_1 + p, i_2 + p, \cdots, i_{2m} + p \pmod{2p}$.

1*st case.* $j \not\equiv i_\sigma + p \pmod{2p}$, $(\sigma = 1, 2, \cdots, 2m)$. We have $\Delta^j{}_{2n-1} \sim R^{n-1}\bar{\delta}_j$, where $\bar{\delta}_j$ is a cycle homologous to δ_j and is in generic position with respect to the cycles δ_i. We may assume that the joints $P_1, P_2, \cdots, P_\gamma$ in (1) are distinct and are not on $\bar{\delta}_j$. Since $\bar{\delta}_j$ does not meet the cycles

[3] By the involutorial property of the $g_n{}^r$ it follows that in the neighborhood of its generic set $n - r$ points of the variable set are uniform functions of the remaining r points.

$\delta_{i_1}, \delta_{i_2}, \cdots, \delta_{i_{2m}}$, the only points of $\Delta_{2r}^{(i)}$ which are on $\Delta^j{}_{2n-1}$ are those obtained by converting one of the α arbitrary points relative to the factor R^α into a semifixed point constrained to remain on $\bar{\delta}_j$. If $\alpha = 0$, i. e., if $m = r$, the two cycles have no points in common and hence

$$(8) \qquad \Delta_{2r}^{i_1 i_2 \cdots i_{2r}} \cdot \Delta^j{}_{2n-1} \sim 0.$$

If $m < r$, then the intersection of the two cycles is necessarily a multiple of the cycle

$$R^{\alpha-1}\delta_{i_1}\delta_{i_2}\cdots\delta_{i_{2m}}\bar{\delta}_j P_1 P_2 \cdots P_\gamma \sim \Delta_{2r-1}^{i_1, i_2, \cdots, i_{2m}, j} \equiv \Delta_{2r-1}^{(i)j}.$$

We now compare the indicatrices of the cycles $\Delta_{2r}^{(i)}$, $\Delta^j{}_{2n-1}$ and $\Delta_{2r-1}^{(i)j}$ at a common generic point $P \equiv (P_1, P_2, \cdots, P_\gamma, \cdots, P_n)$, where $P_{\gamma+1}, \cdots, P_{\gamma+\alpha-1}$ are arbitrary points, $P_{\gamma+\alpha}$ is on $\bar{\delta}_j$ and $P_{\gamma+\alpha+\sigma}$ is on δ_{i_σ}. Let (ξ_i, η_i) be an indicatrix of R at the point P_i. We assume that $\xi_{\gamma+\alpha}, \xi_{\gamma+\alpha+1}, \cdots, \xi_n$ coincide with the positive directions of $\bar{\delta}_j, \delta_{i_1}, \cdots, \delta_{i_{2m}}$ respectively. We have then

Indicatrix $\Delta_{2r}^{(i)}$: $(\xi_{\gamma+1}, \eta_{\gamma+1}, \cdots, \xi_{\gamma+\alpha}, \eta_{\gamma+\alpha}, \xi_{\gamma+\alpha+1}, \cdots, \xi_n)$;
Indicatrix $\Delta^j{}_{2n-1}$: $(\xi_1, \eta_1, \cdots, \xi_{\gamma+\alpha-1}, \eta_{\gamma+\alpha-1}, \xi_{\gamma+\alpha}, \xi_{\gamma+\alpha+1}, \eta_{\gamma+\alpha+1}, \cdots, \xi_n, \eta_n)$;
Indicatrix $\Delta_{2r-1}^{(i)j}$: $(\xi_{\gamma+1}, \eta_{\gamma+1}, \cdots, \xi_{\gamma+\alpha-1}, \eta_{\gamma+\alpha-1}, \xi_{\gamma+\alpha+1}, \cdots, \xi_n, \xi_{\gamma+\alpha})$;
Indicatrix R^n : $(\xi_1, \eta_1, \cdots, \xi_n, \eta_n)$,

or also

Indicatrix $\Delta_{2r}^{(i)}$: (Indicatrix $\Delta_{2r-1}^{(i)j}, \eta_{\gamma+\alpha}$);
Indicatrix $\Delta^j{}_{2n-1}$: $(-1)^m$(Indicatrix $\Delta_{2r-1}^{(i)j}, \xi_1, \eta_1, \cdots, \xi_\gamma, \eta_\gamma, \eta_{\gamma+\alpha+1}, \cdots, \eta_n$);
Indicatrix R^n : $(-1)^m$(Indicatrix $\Delta_{2r-1}^{(i)j}, \eta_{\gamma+\alpha}, \xi_1, \eta_1, \cdots, \xi_\gamma, \eta_\gamma, \eta_{\gamma+\alpha+1}, \cdots, \eta_n$),

and consequently

$$(9) \qquad \Delta_{2r}^{i_1, i_2, \cdots, i_{2m}} \cdot \Delta^j{}_{2n-1} \sim + \Delta_{2r-1}^{i_1, i_2, \cdots, i_{2m}, j}, \qquad m < r.$$

Remark. If j equals one of the indices $i_1, \cdots, i_{2m}$, then the intersection of $\Delta_{2r}^{(i)}$ and $\Delta^j{}_{2n-1}$ is ~ 0. In fact, if for instance $j = i_1$, then the interchange of the factors δ_{i_1} and $\bar{\delta}_j$ changes $\Delta_{2r-1}^{(i)j}$ into $-\Delta_{2r-1}^{(i)j}$. On the other hand the new cycle must be homologous to $\Delta_{2r-1}^{(i)j}$, since we have only replaced δ_{i_1} and $\bar{\delta}_j$ by the homologous cycles $\bar{\delta}_j$ and δ_{i_1}. Hence $\Delta_{2r-1}^{(i)j} \sim -\Delta_{2r-1}^{(i)j}$, and consequently $\Delta_{2r-1}^{(i)j} \sim 0$.[4]

[4] One can prove the homology $\Delta_{2r-1}^{(i)i_1} \sim 0$ also directly by observing that if C_2 is a 2-chain on R bounded by $\bar{\delta}_{i_1} - \delta_{i_1}$, then $R^\alpha \delta_{i_1} \cdots \delta_{i_{2m}} C_2 \to \Delta_{2r-1}^{(i)i_1}$.

2nd case. Let j coincide with one of the numbers $i_\sigma + p \pmod{2p}$, say, let $j \equiv i_1 + p$. In addition to the common points which were present in the 1st case, the two cycles $\Delta_{2r}^{(i)}$ and $\Delta^j{}_{2n-1}$ have now in common also the points obtained by letting the semifixed point on δ_{i_1} coincide with the common point $P_{\gamma+1}$ of δ_{i_1} and δ_{i_1+p}. Hence the complete intersection of the two cycles consists now, in addition to the cycle $\Delta_{2r-1}^{i_1, \cdots, i_{2m}, i_1+p}$ (not present if $m = r$), also of a multiple of the cycle

$$R^a \delta_{i_2} \cdots \delta_{i_{2m}} P_1 \cdots P_\gamma P_{\gamma+1} \sim \Delta_{2r-1}^{i_2, \cdots, i_{2m}}.$$

We have therefore

$$\Delta_{2r}^{i_1, \cdots, i_{2m}} \cdot \Delta_{2n-1}^{i_1+p} \sim \epsilon\, \Delta_{2r-1}^{i_1, \cdots, i_{2m}, i_1+p} + \eta\, \Delta_{2r-1}^{i_2, \cdots, i_{2m}},$$

where, as in the 1st case, we find $\epsilon = 1$, if $m < r$, and $\epsilon = 0$ if $m = r$. To find η we construct indicatrices at a generic point P of $\Delta_{2r-1}^{i_2, \cdots, i_{2m}}$. Let $P \equiv (P_1, \cdots, P_\gamma, P_{\gamma+1}, \cdots, P_n)$, where $P_{\gamma+2}, P_{\gamma+3}, \cdots, P_{\gamma+2m}$ are on $\delta_{i_2}, \delta_{i_3}, \cdots, \delta_{i_{2m}}$ respectively and where $P_{\gamma+2m+1}, \cdots, P_n$ are arbitrary points. We fix an indicatrix (ξ_i, η_i) of R at each point P_i, $i \neq \gamma + 1$. As for $P_{\gamma+1}$, the common point of δ_{i_1} and δ_{i_1+p}, we take for $\xi_{\gamma+1}$ and $\eta_{\gamma+1}$ the positive directions on δ_{i_1} and δ_{i_1+p} respectively, so that at $P_{\gamma+1}$ the indicatrix is $(\delta_{i_1} \cdot \delta_{i_1+p})(\xi_{\gamma+1}, \eta_{\gamma+1})$. We have

Indicatrix $\Delta_{2r}^{i_1, \cdots, i_{2m}}$: $(\xi_{\gamma+1}, \cdots, \xi_{\gamma+2m}, \xi_{\gamma+2m+1}, \eta_{\gamma+2m+1}, \cdots, \xi_n, \eta_n)$;
Indicatrix $\Delta_{2n-1}^{i_1+p}$: $(\eta_{\gamma+1}, \xi_1, \eta_1, \cdots; \xi_\gamma, \eta_\gamma, \xi_{\gamma+2}, \eta_{\gamma+2}, \cdots, \xi_n, \eta_n)$:
Indicatrix $\Delta_{2r-1}^{i_2, \cdots, i_{2m}}$: $(\xi_{\gamma+2}, \cdots, \xi_{\gamma+2m}, \xi_{\gamma+2m+1}, \eta_{\gamma+2m+1}, \cdots, \xi_n, \eta_n)$;
Indicatrix R^n : $(\delta_{i_1} \cdot \delta_{i_1+p})(\xi_1, \eta_1, \cdots, \xi_n, \eta_n)$;
or
Indicatrix $\Delta_{2r}^{i_1, \cdots, i_{2m}}$: $-(\text{Indicatrix } \Delta_{2r-1}^{i_2, \cdots, i_{2m}}, \xi_{\gamma+1})$;
Indicatrix $\Delta_{2n-1}^{i_1+p}$: $(-1)^m(\text{Indicatrix } \Delta_{2r-1}^{i_2, \cdots, i_{2m}}, \eta_{\gamma+1}, \cdots, \eta_{\gamma+2m}, \xi_1, \eta_1, \cdots, \xi_\gamma, \eta_\gamma)$;
Indicatrix R^n : $(-1)^{m-1}(\delta_{i_1} \cdot \delta_{i_1+p})$
$(\text{Indicatrix } \Delta_{2r-1}^{i_2, \cdots, i_{2m}}, \xi_{\gamma+1}, \eta_{\gamma+1}, \eta_{\gamma+2}, \cdots, \eta_{\gamma+2m}, \xi_1, \eta_1, \cdots, \xi_\gamma, \eta_\gamma)$,

and this proves that $\eta = (\delta_{i_1} \cdot \delta_{i_1+p})$. We have thus the following homologies:

$$(10) \quad \Delta_{2r}^{i_1, \cdots, i_{2m}} \cdot \Delta_{2n-1}^{i_1+p} \sim \Delta_{2r-1}^{i_1, \cdots, i_{2m}, i_1+p} + (\delta_{i_1} \cdot \delta_{i_1+p})\Delta_{2r-1}^{i_2, \cdots, i_{2m}}, \quad m < r;$$

$$(11) \quad \Delta_{2r}^{i_1, \cdots, i_{2r}} \cdot \Delta_{2n-1}^{i_1+p} \sim (\delta_{i_1} \cdot \delta_{i_1+p})\Delta_{2r-1}^{i_2, \cdots, i_{2r}}.$$

Before we apply the homologies (8-11) toward the computation of the coefficients $\epsilon_{(i)}$ in (7), we show that already the form of these homologies

permits us to derive a fundamental property of linear series on an algebraic curve. The coefficients $\epsilon_{(i)}$ do not all vanish ($\epsilon = +1$). Let us assume that there is at least one coefficient with 2σ indices which is different from zero ($0 \leqq \sigma \leqq \rho$), say, let $\epsilon_{j_1, \ldots, j_{2\sigma}} \neq 0$, and that all the coefficients $\epsilon_{(i)}$ with more than 2σ indices vanish. We observe that if $m < r$ then in the expression of $\Delta_{2r}^{i_1, \ldots, i_{2m}} \cdot \Delta^j_{2n-1}$ there occurs the cycle $\Delta_{2r-1}^{i_1, \ldots, i_{2m}, j}$ having one more index than $\Delta_{2r}^{i_1, \ldots, i_{2m}}$. Hence, *if we assume that* $\sigma < r$, then we find in the homology (7′) the term $\epsilon_{j_1, \ldots, j_{2\sigma}} \Delta_{2r-1}^{j_1, \ldots, j_{2\sigma}, j}$, and this term obviously does not cancel with any other term. *If, in addition, we assume that* $\sigma < p$, we may choose for j a value distinct from $j_1, j_2, \cdots, j_{2\sigma}$, so that $\Delta_{2r-1}^{j_1, \ldots, j_{2\sigma}, j}$, having distinct indices, is one of the cycles of the minimal base. We thus arrive at a contradiction, since such a cycle cannot occur in the homology (7′) with a coefficient $\neq 0$. It follows that our two assumptions: $\sigma < r$ and $\sigma < p$, cannot be true simultaneously, i. e., *of the two inequalities* $\sigma \geqq r$, $\sigma \geqq p$, *one at least must be true.* Since $\sigma \leqq \rho$ and since ρ is the smallest of the numbers r, p, $n - r$, we conclude that necessarily

$$\sigma = \rho = r \text{ or } p, \tag{12}$$

and, moreover, that *the order n and the dimension r of a linear series $g_n{}^r$ on an algebraic curve of genus p satisfy necessarily at least one of the two inequalities:* $n - r \geqq r$, $n - r \geqq p$. The so-called *special series* are those for which the second inequality does not hold, i. e., those for which $r > n - p$. Consequently we may state the above result as follows: *the order n and the dimension r of a special series $g_n{}^r$ always satisfy the inequality* $n \geqq 2r$. This is the *theorem of Clifford* in its usual formulation. Our recognition of the topological character of this classical theorem is well in agreement with the fact that it is by no means an existence theorem, since it gives only an upper limit for the dimension r of a linear series $g_n{}^r$. One cannot expect topological proofs of existence theorems in algebraic geometry!

4. *Computation of the coefficients* $\epsilon_{(i)}$. Using the intersection formulas (8-11) and taking into account (12) and the remark after formula (9), we find for $\Gamma_{2r} \cdot \Delta^{j+p}_{2n-1}$ the following expression:

$$\begin{aligned} \Gamma_{2r} \cdot \Delta^{j+p}_{2n-1} \sim & \sum_{m=0}^{\rho-1} \sum\nolimits_{(i)} \epsilon_{i_1, i_2, \ldots, i_{2m}} \Delta_{2r-1}^{i_1, i_2, \ldots, i_{2m}, j+p} \\ & + (\delta_j \cdot \delta_{j+p}) \sum_{m=1}^{\rho} \sum\nolimits_{(i)} \epsilon_{j, i_2, \ldots, i_{2m}} \Delta^{i_2, \ldots, i_{2m}} \sim 0. \end{aligned} \tag{13}$$

The first summation is extended to all unordered sets of indices $i_1, \cdots, i_{2m}$

different from $j + p$ (mod $2p$), while the second summation is extended to all unordered sets of indices $i_2, \cdots, i_m$ different from j. Any cycle $\Delta_{2r-1}^{j_1, j_2, \cdots, j_{2m-1}}$ in which the indices are all different from $j + p$ (mod $2p$) occurs in (13) with the coefficient $(\delta_j \cdot \delta_{j+p}) \epsilon_{j, j_1, j_2, \ldots, j_{2m-1}}$. Since this holds for $j = 1, 2, \cdots, 2p$, it follows that if the indices are arranged in order of magnitude then the coefficients $\epsilon_{i_1, i_2, \ldots, i_{2m}}$ which are not of the type $\epsilon_{i_1, \ldots, i_m, i_1+p, \ldots, i_m+p}$ are all zero. If we now consider a cycle $\Delta_{2r-1}^{j_1, \cdots, j_{m-1}, j_1+p, \cdots, j_{m-1}+p, j+p}$, we see that its coefficient in (13) equals

$$\epsilon_{j_1, \ldots, j_{m-1}, j_1+p, \ldots, j_{m-1}+p} + (\delta_j \cdot \delta_{j+p}) \epsilon_{j, j_1, \ldots, j_{m-1}, j_1+p, \ldots, j_{m-1}+p, j+p}.$$

Hence

$$\epsilon_{i_1, \ldots, i_m, i_1+p, \ldots, i_m+p} = (-1)^m (\delta_{i_1} \cdot \delta_{i_1+p}) \epsilon_{i_2, \ldots, i_m, i_2+p, \ldots, i_m+p}.$$

Applying this recurrence relation m times and recalling that $\epsilon = 1$, we obtain

$$\epsilon_{i_1, i_2, \ldots, i_m, i_1+p, i_2+p, \ldots, i_m+p} = (-1)^{[m(m+1)/2]} (\delta_{i_1} \cdot \delta_{i_1+p}) (\delta_{i_2} \cdot \delta_{i_2+p}) \cdots (\delta_{i_m} \cdot \delta_{i_m+p}),$$

or, $\epsilon_{i_1, i_1+p, \ldots, i_m, i_m+p} = (-1)^m$, provided $i_1, \cdots, i_m$ are all less than p. *We have therefore the following expression for the cycle Γ_{2r} associated with the g_n^r:*

$$(14) \qquad \Gamma_{2r} \sim \Delta_{2r} - \Sigma \Delta_{2r}^{i_1, i_1+p} + \Sigma \Delta^{i_1, i_1+p, i_2, i_2+p} + \cdots + (-1)^\rho \Sigma \Delta^{i_1, i_1+p, \cdots, i_\rho, i_\rho+p},$$

where the summation indices $i_1, i_2, \cdots$ are less than p and where $\rho = r$ or p according as $r \leqq p$ or $r \geqq p$.

COROLLARY. *The various linear series g_n^r of a given order n and of a given dimension r on a curve f are represented on R^n by homologous cycles (in the sense of homologies with allowed division).*[5]

5. *The Riemann-Roch theorem.* The following are existence theorems and therefore essentially algebraic in nature:

(a) *There exist infinitely many series g_n^r, of increasing orders, such that $r \geqq n - p$.*

(b) *There exists a series g_{2p-2}^{p-1}.*

These theorems follow in an elementary manner from the consideration

[5] It would be desirable to find out whether R^n does or does not possess torsion, in order to conclude as to the validity of this corollary with respect to homologies without division. It is not true, however, that the linear series g_n^r on f of given order and dimension necessarily form an irreducible algebraic system. This is true only for series of a sufficiently general type, for instance of a sufficiently high order.

of the series cut out on a plane algebraic curve of order m by its adjoint curves of order $l \geqq m-3$. It is found that if $l > m-3$, then the series is of order $2p-2+m(l-m+3)$ and of dimension $\geqq p-2+(l-m+3)$, and if $l=m-3$, then it is of order $2p-2$ and of dimension $\geqq p-1$. We regard theorems (a) and (b) as the algebro-geometric constituents of the Riemann-Roch theorem and we proceed to prove the rest of this theorem topologically. We naturally assume that the topological significance of the two-fold of the genus of an algebraic curve has been already established (for instance, by means of the consideration of the Euler-Poincaré characteristic of the m-sheeted Riemann surface of the curve f). We assume moreover the following two properties of *complete* linear series [6] which follow directly from the definition of linear series by means of linear systems of rational functions on the curve f: 1) two *distinct* complete series of the same order have no sets in common: 2) if a set G of a given complete series $g_{n_2}^{r_2}$ is contained in one or more sets of another complete series $g_{n_1}^{r_1}$ $(n_1 > n_2)$, then the *residual sets* form a *complete* linear series of order $n_1 - n_2$, and this series remains the same as G varies in $g_{n_2}^{r_2}$ (*the residue theorem in its invariantive form*).

(1) *If g_n^r is a special series, then necessarily $n \leqq 2p-2$.* This is a consequence of the theorem of Clifford. In fact, if the g_n^r is special, then $n \geqq 2r \geqq 2(n-p+1)$, i. e., $n \leqq 2p-2$.

From (1) it follows that the series g_n^r of theorem (a) are *complete* and of dimension $r=n-p$, provided $n > 2p-2$. From this and from the fact that these series have an arbitrarily high dimension it follows immediately, by the residue theorem, that

(2) *if g_n^r is any complete series, then $r \geqq n-p$. If $n > 2p-2$, then $r=n-p$.*

For series of order $2p-2$ we now prove the following:

(3) *A series g_{2p-2}^{p-1} is necessarily complete. There cannot exist two distinct series g_{2p-2}^{p-1}.*

The first part of the theorem follows immediately from the theorem of Clifford. To prove the second part of the theorem let us assume that there exist two distinct series g_{2p-2}^{p-1}, and let $\Gamma_{2p-2}^{(1)}$ and $\Gamma_{2p-2}^{(2)}$ be the corresponding cycles on the symmetric product R^{2p-2}. We have from (14):

$$(15) \qquad \Gamma_{2p-2}^{(1)} \sim \Gamma_{2p-2}^{(2)} \sim \Delta_{2p-2} - \Sigma\Delta_{2p-2}^{i_1,\, i_1+p} + \cdots + (-1)^{p-1}\Sigma\Delta_{2p-2}^{i_1,\, i_1+p,\, \ldots,\, i_{p-1},\, i_{p-1}+p}.$$

[6] I. e., series g_n^r which are not contained in linear series of the same order n and of a higher dimension.

Using the intersection formulas (4), (5) we find:

$$(16)\quad (\Gamma^{(1)}_{2p-2}\cdot\Gamma^{(2)}_{2p-2}) = 1 - \binom{p}{1} + \binom{p}{2} + \cdots + (-1)^{p-1}\binom{p}{p-1} = (-1)^{p-1}.$$

The intersection number of the two cycles being different from zero, it follows that the two complete series g^{p-1}_{2p-2} have at least one set in common, and this contradicts the assumption that the two series are distinct, q. e. d.

We have thus proved *the uniqueness of the canonical series* g^{p-1}_{2p-2} and hence also *its invariance under birational transformations.* We observe that by (16) *the virtual degree of the canonical series (as a cycle on* R^{2p-2}*) equals* $(-1)^{p-1}$.

(4) *Any special series* $g_n{}^r$ *is partially contained in the canonical series.*[7]

Proof. By (1) we have necessarily $n \leqq 2p-2$. The case $n = 2p-2$ was already settled in (3), so that we may assume, if we wish, $n < 2p-2$, although the proof below does not require this assumption. It is sufficient to prove the theorem for a series $g_n{}^{n-p+1}$ contained in the given $g_n{}^r$. We consider of the curve f the series of all sets of $2p-2$ points which are made up of a variable set of the $g_n{}^{n-p+1}$ and of a variable set of $2p-2-n$ arbitrary points. This series, which we shall denote by s^{p-1}_{2p-2} is of dimension $p-1$ and is the locus of the linear series $g^{n-p+1}_{2p-2} = g_n{}^{n-p+1} + P_1 + \cdots + P_{2p-2-n}$ as the $2p-2-n$ fixed points P_i of this series vary arbitrarily. We wish to find the cycle Γ_{2p-2} which corresponds to this series on the symmetric product R^{2p-2}, i. e., the expression of this cycle in terms of the basic cycles $\Delta^{(i)}_{2p-2}$. In order not to complicate the notations, we replace the series $g_n{}^{n-p+1}$ by an arbitrary series $g_n{}^r$, we add to the sets of this series k fixed points $P_1, \cdots, P_k$, and we look for the cycle Γ_{2r+2k} which corresponds on R^{n+k} to the algebraic series s^{r+k}_{n+k}, locus of the series $g^r{}_{n+k} = g_n{}^r + P_1 + \cdots + P_k$ as the fixed points P_i of this series vary arbitrarily. Let Γ_{2r} be the cycle which corresponds on R^{n+k} to the series $g^r{}_{n+k}$ and let V_n be the subvariety $R^nP_1P_2\cdots P_k$ of R^{n+k}. As the points P_i vary, the variety V_n varies in an algebraic system $\{V_n\}$ of dimension k. For any preassigned V_n in this system we may assume that the basic cycles $\Delta^{(i)}_{2r}$ of R^{n+k} lie on it, since in the expression of these cycles as symmetric products there occur at least k fixed points. To find the locus of any cycle $\Delta^{(i)}_{2r}$ as the carrying variety V_n varies in $\{V_n\}$, it is only necessary to convert the fixed points $P_1, \cdots, P_k$ into arbitrary points, the effect being that of converting the cycle $\Delta^{(i)}_{2r}$ into the cycle $\Delta^{(i)}_{2r+2k}$ *having the same indices.* However,

[7] I. e., any set of the $g_n{}^r$ is contained in one or more sets of the canonical series. By the residue theorem there exists then the *residual series* g^{ρ}_{2p-2-n} of the canonical series with respect to the $g_n{}^r$, where $\rho + 1$ is the number of linearly independent canonical sets containing a given set of the $g_n{}^r$.

the actual locus of $\Delta_{2r}^{i_1,\cdots,i_{2m}}$ is $\binom{r+k-m}{k}\Delta_{2r+2k}^{i_1,\cdots,i_{2m}}$, since in the expression of the cycle $\Delta_{2r+2k}^{i_1,\cdots,i_{2m}}$ as a symmetric product there occur $r+k-m$ arbitrary points (i. e. the factor R^{r+k-m}), and any k of these points can be identified with the variable points $P_1,\cdots,P_k$, so that each cell of $\Delta_{2r+2k}^{i_1,\cdots,i_{2m}}$ must be counted $\binom{r+k-m}{k}$ times. Consequently, denoting by Γ^*_{2r} the right-hand member of (14) and by Γ^*_{2r+2k} the locus of the cycle Γ^*_{2r}, we have:

$$(17)\qquad \Gamma^*_{2r+2k}=\binom{r+k}{k}\Delta_{2r+2k}-\binom{r+k-1}{k}\Sigma\Delta_{2r+2k}^{i_1,i_1+p} +\cdots+(-1)\binom{r+k-\rho}{k}\Sigma\Delta_{2r+2k}^{i_1,i_1+p,\cdots,i_\rho,i_\rho+p},$$

where $\rho=r$ or p, according as $r\leqq p$ or $r\geqq p$.

Now in the present case it is not difficult to show that *the loci* Γ_{2r+2k}, Γ^*_{2r+2k} *of the homologous cycles* Γ_{2r}, Γ^*_{2r} *are also homologous cycles.* In fact, between any two varieties of the system $\{V_n\}$, say

$$V_n^{(1)}=R^nP_1^{(1)}\cdots P_k^{(1)},\qquad V_n^{(2)}=R^nP_1^{(2)}\cdots P_k^{(2)},$$

there is the following uniquely determined homeomorphism:

$$(O_1,\cdots,O_n,P_1^{(1)},\cdots,P_k^{(1)})\leftrightarrow(O_1,\cdots,O_n,P_1^{(2)},\cdots,P_k^{(2)}),$$

and this homeomorphism reduces to the identity if $V_n^{(1)}$ coincides with $V_n^{(2)}$. Since the elements V_n of the system $\{V_m\}$ are in (1, 1) correspondence with the points of R^k, it follows that Γ_{2r+2k} and Γ^*_{2r+2k} are singular images on R^{n+k} of the direct products $\Gamma_{2r}\times R^k$ and $\Gamma^*_{2r}\times R^k$. On the other hand, if C_{2r+1} is a chain on V_n bounded by $\Gamma_{2r}-\Gamma^*_{2r}$, then the locus of C_{2r+1} is a singular image of the direct product $C_{2r+1}\times R^k$. Hence locus $C_{2r+1}\rightarrow\Gamma_{2r+2k}-\Gamma^*_{2r+2k}$, and consequently $\Gamma_{2r+2k}\sim\Gamma^*_{2r+2k}$, so that the desired expression of the cycle Γ_{2r+2k} is given by the right-hand member of the homology (17).

In the particular case of the algebraic series s_{2p-2}^{p-1} considered above ($r=n-p+1$, $k=2p-2-n=p-1-r$), we find for the corresponding cycle Γ_{2p-2}^{p-1} the following homology:

$$\Gamma_{2p-2}\sim\binom{p-1}{p-1-r}\Delta_{2p-2}-\binom{p-2}{p-1-r}\Sigma\Delta_{2p-2}^{i_1,i_1+p} +\cdots+(-1)^r\binom{p-1-r}{p-1-r}\Sigma\Delta_{2p-2}^{i_1,i_1+p,\cdots,i_r,i_r+p},$$

where now $\rho=r$, since $n\leqq 2p-2$ and hence $r\leqq p-1$ by the theorem of Clifford.

Let Γ'_{2p-2} be the cycle on R^{2p-2} which corresponds to the canonical series:

$$\Gamma'_{2p-2}\sim\Delta_{2p-2}-\Sigma\Delta_{2p-2}^{i_1,i_1+p} +\cdots+(-1)^{p-1}\Sigma\Delta_{2p-2}^{i_1,i_1+p,\cdots,i_{p-1},i_{p-1}+p}.$$

We have

$$(\Gamma_{2p-2}\cdot\Gamma'_{2p-2})=\binom{p-1}{p-1-r}-\binom{p-2}{p-1-r}\binom{p}{1}+\binom{p-3}{p-1-r}\binom{p}{2}+\cdots+(-1)^r\binom{p}{r}=(-1)^r.$$

Proof. Let $\phi(p,r)$ denote the left-hand member of the above identity. We have $\phi(p,p-1)=1-\binom{p}{1}+\binom{p}{2}+\cdots+(-1)^{p-1}\binom{p}{p-1}=(-1)^{p-1}$. Moreover, we have for any $r<p-1$,

$$\begin{aligned}\phi(p,r)&+\phi(p,r+1)\\&=\binom{p}{p-1-r}-\binom{p-1}{p-1-r}\binom{p}{1}+\binom{p-2}{p-1-r}\binom{p}{2}+\cdots+(-1)^{r+1}\binom{p}{r+1}\\&=\binom{p}{p-1-r}\left[1-\binom{r+1}{1}+\binom{r+1}{2}-\cdots+(-1)^{r+1}\right]=0,\quad\text{q. e. d.}\end{aligned}$$

The intersection number $(\Gamma_{2p-2}\cdot\Gamma'_{2p-2})$ being different from zero, it follows that the series s^{p-1}_{2p-2} and the canonical series have at least one set in common, i. e., there exists a set G_n in the $g_n{}^r$ and there exists a set of $2p-2-n$ points P_i such that $G_n+P_1+\cdots+P_{2p-2-n}$ is a canonical set, q. e. d.

(5) The above result is essentially equivalent with the Riemann-Roch theorem. For the convenience of the reader we complete the proof. Let $g_n{}^r$ be a complete special series, $r=n-p+i$, $i>0$. Adding to the sets of this series $i-1$ arbitrary fixed points, we obtain a series $g^r{}_{n+i-1}$ which is still special, and hence, by (4), there exists at least one canonical set which contains a set of the $g_n{}^r$ *and* $i-1$ *arbitrary preassigned points.* As a consequence, *if* j *denotes the number of linearly independent canonical sets containing a given set of the* $g_n{}^r$, *then* $j\geqq i$. On the other hand, the residual series of the canonical series with respect to the $g_n{}^r$ is a g^{j-1}_{2p-2-n}, i. e., a series $g_{n_1}{}^{n_1-p+i_1}$, where $n_1=2p-2-n$ and $i_1=(r+1)+(j-i)>0$. Denoting by j_1 the number of linearly independent canonical sets containing a given set of this residual series, we find as above $j_1\geqq i_1$. Since $j_1=r+1$, it follows $r+1\geqq r+1+(j-i)$, i. e., $j\leqq i$, and consequently $j=i$, q. e. d.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, NEW JERSEY.

REFERENCES.

1. S. Lefschetz, *L'analysis situs et la géometrie algébrique*, Paris, 1924.
2. S. Lefschetz, "Topology," *American Mathematical Society, Colloquium Publications*, vol. 12, New York, 1930.
3. S. Lefschetz, "Correspondences between algebraic curves," *Annals of Mathematics* (2), vol. 28 (1927).
4. M. Richardson, "On the homology characters of symmetric products," *Duke Mathematical Journal*, vol. 1 (1935).
5. J. A. Todd, "Algebraic correspondences between algebraic varieties," *Annals of Mathematics* (2), vol. 36 (1935).
6. O. Zariski, "Algebraic surfaces," *Ergebnisse der Mathematik and ihrer Grenzgebiete*, vol. III, 5, Berlin, Springer, 1935.

COLLINEATION GROUPS IN A FINITE SPACE WITH A LINEAR AND A QUADRATIC INVARIANT.

By Arthur B. Coble.

Introduction. In the theory of types of regular Cremona transformations in S_k determined by n points it appears [cf. [4], pp. 39-41] that these types are determined by the elements of a linear group with integer coefficients, which is generated by the permutations of $x_1, \cdots, x_n$ and the additional element,

$$(1) \qquad A_{1,\ldots,k+1} : \begin{array}{ll} x'_i = x_i + M & (i = 0, 1, \cdots, k+1), \\ x'_j = x_j & (j = k+2, \cdots, n) \\ M = (k-1)x_0 - x_1 - \cdots - x_{k+1}. & \end{array}$$

This group has a quadratic, and a linear, invariant,

$$(2) \qquad \begin{aligned} Q &= (k-1)x_0^2 - x_1^2 \cdots - x_n^2, \\ L &= (k+1)x_0 - x_1 \cdots - x_n. \end{aligned}$$

It is ordinarily of infinite order.

As generators of this group we may take the $(n-1)$ transpositions, (x_1x_2), $(x_2x_3), \cdots, (x_{n-1}x_n)$, and $A_{1,\ldots,k+1}$. All of these n generators are of period two, and they lie in a conjugate set which ordinarily contains an infinite number of elements. These generating involutions are of the simplest type, harmonic perspectivities with a center q and an S_{n-1} of fixed points, the polar S_{n-1} of q as to Q. Furthermore the centers q of these involutions are *points on L*.

If the elements of this group are reduced mod. p a finite group, $\Gamma(p)$, is obtained, which for prime p can be regarded as a group in a finite space with a quadratic and a linear invariant. The elements which reduce to the identity mod. p constitute an invariant subgroup of the original group whose factor group is $\Gamma(p)$. Thus the structure of the original group is dependent upon that of $\Gamma(p)$.

The groups of linear transformations in a finite field, G. F. $(s = p^m)$, with a quadratic invariant have been studied by Dickson ([1], Chaps. VII, VIII) who, in the determination of their structure, has been led to important series of simple groups. The nature of $\Gamma(p)$ above is dependent however upon its involutorial generators which play no part in Dickson's treatment. We are thus led to consider the nature of the collineation groups with a quadratic

invariant which are generated by these involutions I_q, and the subgroups which arise when the additional linear invariant is introduced. For the most difficult phase of the argument, namely, the simplicity of certain subgroups, we naturally depend upon Dickson, but the general course of the argument is quite distinct from that employed in the *Linear Groups*.

In § 1 certain normal forms of Q appropriate to the geometric treatment are derived which facilitate useful enumerations. In § 2 the collineation groups of Q in the finite S_n are discussed, and their constitution is determined [cf. (10), (13), (16)]. In § 3 the corresponding groups for a quadratic and a linear form, generated by involutions I_q for points q on L, are obtained [cf. (16), (17), (18)].

We make no attempt here to apply these conclusions to the group $\Gamma(p)$ obtained from the Cremona types. Since the groups $\Gamma(2)$ have already been discussed by the author,[5] we consider only the case of an odd prime.

1. Types of proper quadrics in S_n. We are concerned with the geometry of the proper quadric in a finite linear projective space S_n of dimension n which is defined in the Galois field, G. F. $[s = p^m]$, p an odd prime. The number of points in such a space is

$$P_n = (s^{n+1} - 1)/(s - 1). \tag{1}$$

For our purposes we shall usually need to distinguish only two classes of such fields: namely, those for which -1 is a square σ, or a not-square ν. These we denote by

$$(2) \quad \begin{array}{llll} \text{F. G. (I)}: & -1 = \sigma; & m \not\equiv 1 \ (\text{mod. } 2) & \text{or} \quad p \not\equiv 3 \ (\text{mod. } 4); \\ \text{F. G. (II)}: & -1 = \nu; & m \equiv 1 \ (\text{mod. } 2) & \text{and} \quad p \equiv 3 \ (\text{mod. } 4). \end{array}$$

In such a space, and with $p > 2$, the usual theory of harmonic pairs is valid. Thus the customary rational reduction of the proper quadratic form whose discriminant is not zero to a sum of $n + 1$ squares is also valid. We may therefore [cf. also [1], § 168] write the quadric in the form

$$Q(n) \equiv \lambda_0 x_0^2 + \lambda_1 x_1^3 + \cdots + \lambda_n x_n^2 \qquad (\lambda_i \neq 0). \tag{3}$$

Obviously $Q(1)$ contains two, or no, real points according as $-\lambda_0\lambda_1$ is a σ, or a ν. Also $Q(2)$ contains $s + 1$ real points [cf. [1], § 64]. Indeed $Q(2)$ is the usual locus generated by two projective line pencils. Any $Q(n)$ $[n > 2]$ has $Q(2)$ sections obtained by setting all but three of the x's in (3) equal to zero, whence

(4) *Every $Q(n)$, $n \geqq 2$, contains real points and proper bisecants.*

Let y, z be two points of $Q(n)$ whose join yz is a proper bisecant. If y, z are chosen as the last two reference points, and if their respective polar spaces are chosen as $x_n = 0$, $x_{n-1} = 0$, then $Q(n)$ takes the form $Q(n-2) + 2\alpha_{n-1,n}x_{n-1}x_n$. The coefficient $2\alpha_{n-1,n}$ may be removed, or changed at will, by a multiplication. This process can be applied to $Q(n-2)$, and continued until a $Q(0)$ is reached if n is even, or until a $Q(1)$ is reached if n is odd. Hence

(5) *Every proper quadratic form in S_n can be reduced by linear transformation with coefficients in the* G. F. (s) *to one of the following forms:*

$$\begin{aligned} n \textit{ even:} \quad & Q(n) = x_0^2 + x_1x_2 + \cdots + x_{n-1}x_n, \\ & \nu Q(n) = \nu(x_0^2 + x_1x_2 + \cdots + x_{n-1}x_n); \\ n \textit{ odd:} \quad & Q_+(n) = x_0x_1 + x_2x_3 + \cdots + x_{n-1}x_n, \\ & Q_-(n) = -\nu x_0^2 + x_1^2 + x_2x_3 + \cdots + x_{n-1}x_n. \end{aligned}$$

Thus in even spaces there is but one geometric type of quadratic, whereas in odd spaces there are two types. We shall see that the notation is so chosen that $Q_+(n)$ is on more points than $Q_-(n)$.

We shall divide the points x of S_n with respect to $Q(n)$ into *outside points, inside points,* and *quadric points* according as $Q(n)(x) = \sigma, \nu, 0$.

We shall find that, when n is even, the number of outside points of $Q(n)$ is greater than the number of inside points. It is for this reason that we have preferred the form $Q(n)$ over $\nu Q(n)$. For this reason also the form $Q(n)$ cannot be linearly transformed into the form $\nu Q(n)$. On the other hand,

(6) *When n is odd, there exist linear transformations which convert $Q_\pm(n)$ into $\nu Q_\pm(n)$ and thus interchange the inside and outside points of $Q_\pm(n)$.*

Indeed, for the type $Q_+(n)$ an obvious linear transformation is $x_{2i} = \nu' x'_{2i}$, $x_{2i+1} = x'_{2i+1}$ $(i = 0, \cdots, (n-1)/2)$. For the type $Q_-(n)$, a similarly formed transformation takes care of the product terms. If $-1 = \sigma$, the additional equations, $x_0 = \{\nu'/\sigma\nu\}^{1/2}x'_1$, $x_1 = \{\sigma\nu\nu'\}^{1/2}x'_0$, take care of the square terms. However, if $-1 = \nu''$, we have to transform $\nu''\nu x_0^2 + x_1^2$ into $\nu''\nu'\nu x'^2_0 + \nu' x'^2_1$. Dickson [1, § 169] shows that $\nu(x_r^2 + x_s^2)$ can be transformed into $x'^2_r + x'^2_s$, whence $\nu''\nu x_0^2 + x_1^2$ can be transformed into $\nu'(\nu''\nu x'^2_0 + x'^2_1)$. When $Q_\pm(n)(x) = \nu' Q_\pm(n)(x')$, due to the linear transformation, then, if $Q_\pm(n)(x) = \sigma$ or ν; $Q_\pm(n)(x') = \sigma/\nu'$ or ν/ν'; i. e., outside and inside points of Q are interchanged.

Thus, when n is odd, the inside and outside points of $Q_\pm(n)$ play the same geometric rôle with respect to $Q_\pm(n)$. When n is even, they do not [cf. (11)].

In obtaining the canonical form (2), the polar space x'_0 of any point p' not on Q gives rise to a term $\lambda'_0 x'^2_0$. This point p' is conjugate to the first reference point under a collineation which leaves Q unaltered if λ_0, λ'_0 are both squares or both not-squares. Thus all the outside points are conjugate, and all the inside points are conjugate, under the collineation group of Q. On applying a similar argument to the term $x_{n-1}x_n$ of the canonical forms (5), we obtain the theorem:

(7) *Under the collineation group of Q, the outside points of Q, the inside points of Q, the pairs of points (in either order) on Q and on a proper bisecant, and the proper bisecants of Q, each form a conjugate set.*

We compare the canonical forms (5) of Q with the types employed by Dickson. These are

$$(8) \qquad \begin{array}{lll} A: & & y_0^2 + y_1^2 + \cdots + y_n^2, \\ B: & n \text{ odd}: & \nu y_0^2 + y_1^2 + \cdots + y_n^2, \end{array}$$

the type A occurring for all values of n. An obvious transformation converts the product $x_i x_{i+1}$ into $x'^2_i - x'^2_{i+1}$, and this in G. F. (I) into $x'^2_i + x''^2_{i+1}$. Thus each product yields two squares with unit coefficients in G. F. (I), whereas in G. F. (II) it yields two squares with one coefficient ν. As noted above, two squares with coefficients ν, ν' can be converted into two squares with unit coefficients. Hence, in G. F. (I), $Q(n)$ yields the type A. In G. F. (II), $Q(n)$ yields $n/2$ coefficients ν and $\nu Q(n)$ yields $1 + n/2$ coefficients ν. Thus $Q(n)$ is of type A if $n/2$ is even and $\nu Q(n)$ is of type A if $n/2$ is odd. In G. F. (I), $Q_+(n)$ is of type A and $Q_-(n)$ is of type B. In G. F. (II), $Q_+(n)$ is of type A or B according as $(n+1)/2$ is even or odd, while $Q_-(n)$ is of type A or B according as $(n-1)/2$ is even or odd. Hence

(9) *When n is even, type A is $Q(n)$ except in* G. F. (II) *for $n \equiv 2$ (mod. 4) when it is $\nu Q(n)$. When n is odd, types A, B are $Q_+(n)$, $Q_-(n)$ respectively except in* G. F. (II) *for $n \equiv 1$ (mod. 4) when they are $Q_-(n)$, $Q_+(n)$ respectively.*

Let $q(n)$, $o(n)$, $i(n)$ denote the number of points in S_n which are respectively *on, outside, inside* the quadric $Q(n)$, this quadric being one of the three types in (5).

We divide the points of S_n into the following four classes:

$$\begin{array}{ll} \text{(a)} & y_0, y_1, \cdots, y_{n-2}, 0, 0; \\ \text{(b)} & y_0, y_1, \cdots, y_{n-2}, 1, 0; \\ & y_0, y_1, \cdots, y_{n-2}, 0, 1; \end{array}$$

(c) $y_0, y_1, \cdots, y_{n-2}, y, 1;$

(d) $0, 0, \cdots, 0, \quad 1, 0;$

$0, 0, \cdots, 0, \quad 0, 1;$

$0, 0, \cdots, 0, \quad y, 1 \qquad (y \neq 0).$

The number of points of type (a) on $Q(n)$ is $q(n-2)$; of each type (b) is $(s-1)$. $q(n-2)$, since a factor $\lambda \neq 0$ must be allowed for in y_i; of type (c) is $\{P_{n-2} - q(n-2)\}(s-1)$, since, for each point not on $Q(n-2)$ with any non-zero multiplier λ, y is unique; and of type (d) is two. Thus we have the recursion formula

$$q(n) = (s-1)P_{n-2} + sq(n-2) + 2,$$

with the initial conditions, $q(0) = 0$, $q_+(1) = 2$, $q_-(1) = 0$. This recursion formula is satisfied by the values given in (10).

According to (6), $o_\pm(n) = i_\pm(n)$. Since also $o_\pm(n) + i_\pm(n) + q_\pm(n) = P_n$, and $q_\pm(n)$ has just been determined, the values $o_\pm(n)$, $i_\pm(n)$ must be those given in (10).

There remains the case $Q(n)$, n even. We ask for $o(n)$, the number of points x for which $Q(n)(x)$ is a square $\sigma \neq 0$. The four classes of points above contribute to this number as follows: (a) $o(n-2)$; (b) $2(s-1)o(n-2)$, allowing for a factor $\lambda \neq 0$ in y_i;

$$\text{(c)} \quad (s-1)\cdot i(n-2)\cdot(s-1)/2 + (s-1)\cdot o(n-2)\cdot(s-3)/2 \\ + (s-1)\cdot q(n-2)\cdot(s-1)/2$$

and (d) $(s-1)/2$. The case (c) needs some explanation. Any $x_0, \cdots, x_{n-2}$ set in Q_{n-2}, after multiplication by $\lambda \neq 0$, yields a ν', σ', 0 in respectively $(s-1)\cdot i(n-2)$, $(s-1)o(n-2)$, $(s-1)q(n-2)$ cases. When ν' occurs we have to pick a $y \neq 0$ for which $\nu' + y = \sigma$. Thus for each of the $(s-1)/2$ squares σ there is a $y \neq 0$. When σ' occurs, in $\sigma' + y = \sigma$ the square $\sigma = \sigma'$ must be avoided to secure $y \neq 0$. On simplifying the total number of points by using $o(n-2) + i(n-2) + q(n-2) = P_{n-2}$, we obtain the recursion formula,

$$o(n) = s\cdot o(n-2) + P_{n-2}\cdot(s-1)^2/2 + (s-1)/2,$$

with the initial value $o(0) = 1$. This yields $o(n)$ in (10), and $i(n)$ in (10) is obtained from $o(n) + i(n) + q(n) = P_n$. Hence the complete enumeration is

$$(10) \quad \begin{array}{ll} n \text{ even: } q(n),\ 2o(n),\ 2i(n) = P_{n-1} & s^n + s^{n/2},\ s^n - s^{n/2}, \\ n \text{ odd: } q_\pm(n), 2o_\pm(n), 2i_\pm(n) = P_{n-1} \pm s^{(n-1)/2}, & s^n \mp s^{(n-1)/2},\ s^n \mp s^{(n-1)/2}. \end{array}$$

In order to enumerate the various types of lines in S_n with respect to $Q(n)$ we need the theorem:

(11) *When n is even, the section of $Q(n)$ by the polar space of an outside [inside] point of $Q(n)$ is a $Q_+(n-1)$ $[Q_-(n-1)]$. When n is odd, the section of $Q_+(n)$ by the polar space of an outside [inside] point of $Q_+(n)$ is a $Q(n-1)$ $[\nu Q(n-1)]$ in* G. F. (I), *but a $\nu Q(n-1)$ $[Q(n-1)]$ in* G. F. (II); *for $Q_-(n)$, the sections $Q(n-1)$ and $\nu Q(n-1)$ are reversed.*

In view of (7) it is sufficient to verify (11) for particular outside and inside points such as $0, \cdots, 0, 1, 1$ and $0, \cdots, 0, 1, \nu$. We observe again the geometric difference between outside and inside points of $Q(n)$, a difference lacking for $Q_\pm(n)$.

A line will be called a *skew line*, a *tangent*, a *secant* (proper bisecant), or a *generator*, of a quadric $Q(n)$ if it meets $Q(n)$ in 0, 1, 2 or more, distinct and real points. Let $\rho_0, \rho_1, \rho_2, \rho_3$ be the number of these respective lines on a point r of S_n.

Since P_{n-1} is the total number of lines on a point of S_n, we have the relations:

$$(\alpha) \qquad \begin{array}{ll} r \text{ on } & Q(n): \rho_1 + \rho_2 + \rho_3 = P_{n-1}, \\ r \text{ not on } & Q(n): \rho_0 + \rho_1 + \rho_2 = P_{n-1}. \end{array}$$

For any type of quadric, and a quadric point, the number of tangents and generators on the point is P_{n-2}, the number of lines on the point and in its polar space; and the number of generators is $q(n-2)$. Thus $\rho_1 + \rho_3 = P_{n-2}$, $\rho_3 = q(n-2)$ together with (α), yields the first enumeration in (12).

If n is odd, outside and inside points are conjugate under linear transformation. For p a point of either type, ρ_1 is the number, $q(n-1)$, of points in the section of $Q_\pm(n)$ by the polar space of p. The remaining $q_\pm(n) - q(n-1)$ points of $Q_\pm(n)$ are paired on ρ_2 secants. These facts, with (α), yield the second enumeration in (12).

If n is even, on an outside point $\rho_1 = q_+(n-1)$, and

$$2\rho_2 = q(n) - q_+(n-1);$$

on an inside point $\rho_1 = q_-(n-1)$, and $2\rho_2 = q(n) - q_-(n-1)$ [cf. (11)]. Thus, with (α), we have the third enumeration in (12). Hence

(12) *On a point of the quadric*

$$\begin{aligned} Q(n): \ & \rho_1, \rho_2, \rho_3 = s^{n-2}, \ s^{n-1}, P_{n-3}; \\ Q_+(n): \ & \rho_1, \rho_2, \rho_3 = s^{n-2} - s^{(n-3)/2}, s^{n-1}, P_{n-3} + s^{(n-3)/2}; \\ Q_-(n): \ & \rho_1, \rho_2, \rho_3 = s^{n-2} + s^{(n-3)/2}, s^{n-1}, P_{n-3} - s^{(n-3)/2}. \end{aligned}$$

For a quadric $Q_{\pm}(n)$, and either an outside or inside point,

$$2\rho_0, 2\rho_1, 2\rho_2 = s^{n-1} \mp s^{(n-1)/2}, 2P_{n-2}, s^{n-1} \pm s^{(n-1)/2}.$$

For a quadric $Q(n)$, and for an

outside point: $2\rho_0, 2\rho_1, 2\rho_2 = s^{n-1} - s^{(n-2)/2}, 2(P_{n-2} + s^{(n-2)/2}), s^{n-1} - s^{(n-2)/2}$;
inside point: $2\rho_0, 2\rho_1, 2\rho_2 = s^{n-1} + s^{(n-2)/2}, 2(P_{n-2} - s^{(n-2)/2}), s^{n-1} + s^{(n-2)/2}$.

We note some further facts employed in the sequel, using the notation p_o and p_i for outside and inside points respectively.

(13) *A secant of Q cuts Q in two points and contains $(s-1)/2$ points p_o and $(s-1)/2$ points p_i. A tangent of Q contains one point of Q and either s points p_o or s points p_i. A skew line of Q contains $(s+1)/2$ points p_o and $(s+1)/2$ points p_i.*

For, if y, z are two points on $Q = (\alpha x)^2$, and yz is a secant, the point $\lambda_1 y + \lambda_2 z$ substituted in Q yields $2\lambda_1\lambda_2(\alpha y)(\alpha z)$. If y, z are not on Q and yz is a skew line of Q, the similar result is $\lambda_1^2(\alpha y)^2 + 2\lambda_1\lambda_2(\alpha y)(\alpha z) + \lambda_2^2(\alpha z)^2$, an imaginary pair in $\lambda_1 : \lambda_2$. In these two cases, for variable $\lambda_1 : \lambda_2$, the theorem follows from (10) for $n = 1$. If y is on Q and z is on a tangent at y we get $\lambda_2^2(\alpha z)^2$ which has the squarity of $(\alpha z)^2$.

(14) *A harmonic pair of $Q_+(1)$ is, in* G. F. (I), *a pair of points p_o, or a pair of points p_i; in* G. F. (II), *a point p_o and a point p_i. A harmonic pair of $Q_-(1)$ is, in* G. F. (I), *a point p_o and a point p_i; in* G. F. (II) *a pair of points p_o or a pair of points p_i.*

For, with $Q_+(1) = x_0x_1$, the harmonic pair $x_0 : x_1$, $-x_1 : x_0$ yields x_0x_1, $-x_0x_1$, which have the same squarity in G. F. (I) but not in G. F. (II). With $Q_-(1) = -\nu x_0^2 + x_1^2$, the harmonic pair $x_0 : x_1$, $x_1 : \nu x_0$ yields $-\nu x_0^2 + x_1^2$, $-\nu x_1^2 + \nu^2 x_0$, which differ by the factor $-\nu$, a not-square in G. F. (I), a square in G. F. (II).

2. The collineation groups of the quadrics $Q(n)$ in S_n. We wish to determine first the order of the group of collineations which leaves the quadric $Q(n)$ in S_n unaltered, i. e., the collineations whose linear transformations leave the form $Q(n)$ invariant to within a factor. We denote this order generically by $N(n)$, or, more specifically, by $N(n)$ [n even]; $N_+(n)$, $N_-(n)$ [n odd].

According to **1** (10), (12) the number of secants of $Q(n)$ is $q(n) \cdot s^{n-1}/2$.

The number of ordered pairs of points of $Q(n)$ on a secant is $q(n)\cdot s^{n-1}$. Two such points being taken as the last two reference points, and their polar spaces as x_n, x_{n-1} respectively, $Q(n)$ takes the form,

$$Q(n-2)+2a_{n-1,n}x_{n-1}x_n.$$

Any collineation which leaves the last two reference points unaltered, and which leaves $Q(n)$ unaltered to within a factor, is made up of one of the $N(n-2)$ collineations which converts $Q(n-2)$ into $\lambda Q(n-2)$, and of $x_{n-1}=\lambda_{n-1}x'_{n-1}$, $x_n=\lambda_n x'_n$, where $\lambda_{n-1}\lambda_n=\lambda$. This furnishes $s-1$ choices for $\lambda_{n-1}\lambda_n$, whence

(1) $$N(n)=(s-1)\cdot s^{n-1}\cdot q(n)\cdot N(n-2).$$

To obtain initial conditions we observe that $Q(2)$, a conic, contains $s+1$ real points, and that a collineation of $Q(2)$ is obtained from three corresponding pairs, whence $N(2)=(s+1)s(s-1)=(s^2-1)s$. A collineation which leaves $Q_+(1)$, two real points on a line, unaltered must leave each point unaltered or must interchange them (two choices), and must send a third point into any one of $s-1$ places, whence $N_+(1)=2(s-1)$. Similarly, in the extended field determined by the imaginary pair $Q_-(1)$, $N_-(1)=2(s+1)$. Thus we have the conditions,

(2) $$N(2)=(s^2-1)s, \qquad N_\pm(1)=2(s\mp 1).$$

For n even, $q(n)\cdot(s-1)=P_{n-1}\cdot(s-1)$ [cf. 1 (10)] $=s^n-1$, whence

$$N(n)=(s^n-1)s^{n-1}(s^{n-2}-1)s^{n-3}\cdots(s^2-1)s.$$

For n odd,

$$q_\pm(n)\cdot(s-1)=\{P_{n-1}\pm s^{(n-1)/2}\}(s-1)=(s^{(n-1)/2}\pm 1)(s^{(n+1)/2}\mp 1).$$

This leads to

$$\begin{aligned}N_\pm(n)&=(s^{(n+1)/2}\mp 1)(s^{(n-1)/2}\pm 1)s^{n-1}\cdot N_\pm(n-2)\\&=(s^{(n+1)/2}\mp 1)(s^{(n-1)/2}\pm 1)s^{n-1}\cdot(s^{(n-1)/2}\mp 1)(s^{(n-3)/2}\pm 1)s^{n-3}\cdot N_\pm(n-4)\\&=(s^{(n+1)/2}\mp 1)s^{n-1}(s^{n-1}-1)s^{n-3}\cdot(s^{(n-3)/2}\pm 1)N_\pm(n-4).\end{aligned}$$

Recalling the values (2) for $N_\pm(1)$, we find that

(3) *The order of the collineation group of the quadric is*

n even: $N(n)=(s^n-1)s^{n-1}(s^{n-2}-1)s^{n-3}\cdots(s^2-1)s;$

n odd: $N_\pm(n)=2(s^{(n+1)/2}\mp 1)s^{n-1}(s^{n-1}-1)s^{n-3}(s^{n-3}-1)\cdots s^2(s^2-1).$

If p is any point not on the quadric $Q=(\alpha x)^2$, we denote by I_p the perspective involution with center p, and linear space of fixed points π, π being the polar space of p as to Q. The equations of this involution are

$$x'=x-p\cdot 2(\alpha p)(\alpha x)/(\alpha p)^2. \tag{4}$$

A set of n linearly independent points in $\pi=(\alpha p)(\alpha x)=0$ are each fixed under (4) with respective multipliers $+1$, and p is fixed with multiplier -1, whence I_p is an involution with determinant -1. For it Q is an absolute invariant, i. e.,

$$(\alpha x')^2=(\alpha x)^2. \tag{5}$$

We seek now to determine the collineation groups generated by these involutions I_p. For this some lemmas are necessary.

(6) *The points of Q are conjugate under sequences of involutions I_p.*

Let a, a' be any two distinct points of Q. If aa' is a secant and p is a further point on this secant, then I_p sends a into a'. If aa' is a generator, let b be a point of Q not on the polar space π_a of a. The points of Q not on π_a do not themselves lie in a linear space since Q is not a pair of such spaces. Hence the polar space $\pi_{a'}$ will not exhaust the points b, i. e., points b exist such that ab, $a'b$ are secants. If p, q are points not on Q but on these respective secants, then I_pI_q sends a into a'.

(7) *The ordered pairs of points on Q on secant lines are conjugate under sequences of involutions I_p.*

Let a, b and a', b' be two such ordered pairs. We first send a into a' by a sequence of I_p's, b then going into b''. The plane $a'b''b'$ meets Q in a conic K since $a'b''$ and $a'b'$ are secants and $a'b''b'$ is a proper triangle. If $b'b''$ is a secant of K which meets the tangent $\pi_{a'}$ in p, then I_p leaves a' unaltered and sends b'' into b'. If $b'b''$ is a generator, K is made up of $b'b''$ and a line on a' which meets $b'b''$ in f. If $\bar{b}$ is a third point on $a'f$, and q a third point on $\bar{b}b'$, such that qa' meets $b''\bar{b}$ at p, then I_pI_q sends a' into itself and sends b'' into b'.

Let now T_n be any collineation which leaves $Q(n)$ unaltered. Let a_{n-1}, a_n be the poles of x_n, x_{n-1} respectively, and let T_n send a_{n-1}, a_n into a'_{n-1}, a'_n. According to (7) there is a product, Π_{n-1}, of involutions I_p which sends a'_{n-1}, a'_n into a_{n-1}, a_n and leaves $Q(n)$ unaltered. Then $T_{n-2}=T_n\Pi_{n-1}$ is a collineation which leaves $Q(n)$, a_{n-1}, a_n each unaltered. Hence T_{n-2} is a

collineation on the variables $x_0, \cdots, x_{n-2}$ alone which leaves $Q(n-2)$ unaltered combined with the multiplication $x_{n-1} = \lambda_{n-1}x'_{n-1}$, $x_n = \lambda_n x'_n$.

If n is even, this process can be continued until a transformation of the form,

$$T_n\Pi(I_p) \,:\, x_0 = \lambda_0 x'_0, x_1 = \lambda_1 x'_1, \cdots, x_n = \lambda_n x'_n,$$

is obtained, where

$$\lambda_0^2 = \lambda_1\lambda_2 = \cdots = \lambda_{n-1}\lambda_n.$$

Dividing through by λ_0, the conditions on the new multipliers are

$$1 = \lambda_1\lambda_2 = \cdots = \lambda_{n-1}\lambda_n.$$

But $x_1 = \lambda_1 x'_1$, $x_2 = \lambda_2 x'_2$ $(\lambda_1\lambda_2 = 1)$ is the product of $x_1 = \lambda_1 x'_2$, $x_2 = \lambda_1^{-1}x'_1$ and $x_1' = x_2''$, $x_2' = x_1''$, the other variables being unaltered. These two factors are involutions I_p, whence $T_n\Pi(I_p) = \Pi'(I_p)$, or $T_n = \Pi''(I_p)$.

In the case n odd for $Q_+(n)$ we can find $\Pi(I_p)$ such that $T_n\Pi(I_p)$ is $x_i = \lambda_i x'_i$ $(i = 0, \cdots, n)$, where

$$\lambda_0\lambda_1 = \lambda_2\lambda_3 = \cdots = \lambda_{n-1}\lambda_n = \mu.$$

Case (a). If $\mu = \epsilon^{2l}$ (ϵ a primitive root in G. F.) and the multipliers be divided by ϵ^l we have as before that $T_n\Pi(I_p) = \Pi'(I_p)$ and $T_n = \Pi''(I_p)$.

Case (b). If $\mu = \epsilon^{2l+1}$, and the multipliers be divided by ϵ^l we get $T_n\Pi(I_p) = \tau\Pi'(I_p)$, where

$$\tau:\quad x_{2i} = \epsilon x'_{2i},\ x_{2i+1} = x'_{2i+1} \qquad [i = 0, \cdots, (n-1)/2].$$

Then $T_n = \tau\Pi''(I_p)$. Since I_p leaves $Q_+(n)$ absolutely unaltered, and τ reproduces it multiplied by ϵ, I_p does not interchange outside and inside points, whereas τ does.

In the case of $Q_-(n)$ we can find $\Pi(I_p)$ such that $T_n\Pi(I_p)$ is

$$x_0 = \alpha x'_0 + \beta x'_1,\ x_1 = \gamma x_0 + \delta x'_0,\ x_i = \lambda_i x'_i \qquad (i = 2, \cdots, n),$$

where

$$\lambda_2\lambda_3 = \lambda_4\lambda_5 = \cdots = \lambda_{n-1}\lambda_n = \mu,$$

and

$$(-\nu x_0^2 + x_1^2) = \mu(-\nu x'^2_0 + x'^2_1).$$

The involutorial elements which carry $-\nu x_0^2 + x_1^2$ into a multiple of itself are the reflections in the members of the pencil (variable ρ) of quadratic forms, $\nu x_0^2 + x_1^2 + \rho x_0 x_1$, apolar to $-\nu x_0^2 + x_1^2$. The discriminant, $\Delta = \rho^2 - 4\nu$, of a member is not zero in G. F. The $s+1$ values $x_0 : x_1$ divide into $(s+1)/2$

real apolar pairs ($\Delta = \sigma$), and the remaining $(s+1)/2$ members have imaginary roots ($\Delta = \nu'$). The polarized quadratic yields the reflection, $x'_0 = \rho x_0 + 2x_1, x'_1 = -2\nu x_0 - \rho x_1$, for which $-\nu x'^2_0 + x'^2_1 = \Delta[-\nu x_0^2 + x_1^2]$. The reflections in the real pairs generate a dihedral collineation g_{s+1}, for which factors of proportionality may be so chosen that $-\nu x_0^2 + x_1^2$ is absolutely unaltered; the reflections in the imaginary pairs change $-\nu x_0^2 + x_1^2$ into $\nu'(-\nu x'^2_0 + x'^2_1)$. Hence

Case (a). If μ is a square, $T_n \Pi(I_p) = \Pi'(I_p)$, or $T_n = \Pi''(I_p)$.

Case (b). If μ is a not-square, $T_n \Pi(I_p) = \tau' \Pi'(I_p)$, where τ' is the collineation in $x_0 : x_1$ above augmented by

$$x_{2i} = \mu x'_{2i},\ x_{2i+1} = x'_{2i+1} \qquad [i = 1, \cdots, (n-1)/2].$$

Hence

(8) *The collineation group generated by involutions I_p are, for n even, the entire collineation group of $Q(n)$ of order $N(n)$; and, for n odd, those invariant subgroups of the collineation groups of $Q_\pm(n)$ of index 2 and orders $N_\pm(n)/2$, which do not interchange the inside and outside points of $Q_\pm(n)$.*

We shall denote these groups generated by involutions I_p by $G(n)(I_p)$, $G_+(n)(I_p)$, $G_-(n)(I_p)$ respectively. We seek now to determine their structure.

An involution I_p effects an even or an odd permutation of the points of Q according as ρ_2 for the point p is even or odd. We examine then the parity of ρ_2 as given in **1** (12) and find that:

(9) *The parity of ρ_2 for a point p_o or a point p_i with respect to Q is given in the table:*

	$Q(n)$			$Q_+(n)$	$Q_-(n)$
(mod. 4)	p_o	p_i	(mod. 4)	p_o, p_i	p_o, p_i
$n \equiv 0$	*even*	*odd*	$n \equiv 1$	*odd*	*even*
$n \equiv 2$			$n = 3$		
G. F. (I)	*even*	*odd*	G. F. (I)	*odd*	*even*
G. F. (II)	*odd*	*even*.	G. F. (II)	*even*	*odd*.

We have seen in (8) that for n even the collineation group of $Q(n)$ is generated by involutions I_p. According to (9) some of these effect odd permutations of the points of $Q(n)$, whence $G(n)(I_p)$ has an invariant subgroup of index 2. The points p for which the I_p are even are points p_o except when $n \equiv 2 \pmod 4$ in G. F. (II). But this exceptional case is precisely that in which $\nu Q(n) = A$, or $Q(n) = \nu' A$ [cf. **1** (9)]. If this factor ν' is removed the points p_i become points p_o. Hence

(10) *The collineation group of the quadric,*

$$A \equiv y_0^2 + y_1^2 + \cdots + y_n^2 \text{ (} n \text{ even)},$$

of order $N(n)$ *is generated by involutions* I_p [cf. (8)]. *It contains a simple invariant subgroup of index two and order* $N(n)/2 = FO(n+1, p^m)$, *which is generated by involutions* I_{po}.

In connection with the proof of this theorem we observe that the notation $FO(n+1, p^m)$ is that of Dickson ([1], p. 191). The simplicity of the invariant subgroup of this order (except in the case $n = 2$, $p^m = 3$, when it is the even G_{12} on the four points of the conic $y_0^2 + y_1^2 + y_2^2$) is proved by Dickson. Dickson's groups of linear transformations of determinant unity which leave Q absolutely unaltered must have series of composition whose indices coincide with those of our collineation groups, except for an index 2 when n is odd due to the factor of proportionality ± 1, or except for indices (factors of $s - 1$) which arise from the fact that our collineations do not leave Q absolutely unaltered. Since none of these exceptional indices could have the value $FO(n+1, p^m)$, the simplicity of the collineation group here obtained must follow. That the group is generated by involutions I_{po} is due to the fact that these involutions must generate a subgroup invariant under the group of order $N(n)$. These considerations apply also in the demonstration of (13) and (16).

For odd n the quadrics $Q_{\pm}(n)$ contain systems of linear spaces $S_{(n-1)/2}$ which in the case of $Q_+(n)$ are real; of $Q_-(n)$, are conjugate imaginary. We recall the theorem of C. Segre [2] [cf. also Bertini [3]]:

(11) *If* n *is odd, a proper quadric in* S_n *contains two systems of linear spaces* $S_{(n-1)/2}$. *If* $n \equiv 1$ *mod.* 4 $\{n \equiv 3$ *mod.* 4$\}$, *two* $S_{(n-1)/2}$'s *belong to the same or different systems according as they have an* S_{2d} $\{S_{2d-1}\}$ *or an* S_{2d-1} $\{S_{2d}\}$ *in common* $[2d, 2d-1 < (n-1)/2]$.

In the case of $Q_+(n)$ one such $S_{(n-1)/2}$ on $Q_+(n)$ is $x_0 = x_2 = \cdots = x_{n-1} = 0$. This is transformed by the I_p which interchanges x_0, x_1 into

$$x_1 = x_2 = x_4 = \cdots = x_{n-1} = 0.$$

These two $S_{(n-1)/2}$'s meet in the $S_{(n-3)/2}$,

$$x_0 = x_1 = x_2 = x_4 = \cdots = x_{n-1} = 0.$$

If $n \equiv 1$ mod. 4, $(n-3)/2 = 2d - 1$. If $n \equiv 3$ *mod.* 4, $(n-3)/2 = 2d$. In either case, the two $S_{(n-1)/2}$'s belong to different systems. By continuous

variation this interchange takes place throughout the systems. Since all points p are conjugate under the group of $Q_+(n)$, each of the involutions I_p interchanges the two systems.

In the case of $Q_-(n)$, the $S_{(n-1)/2}$ given by

$$\sqrt{\nu}\, x_0 + x_1 = x_2 = x_4 = \cdots = x_{n-1} = 0$$

is transformed by the I_p which changes the sign of x_0 into its conjugate imaginary. These two $S_{(n-1)/2}$'s meet in the same $S_{(n-3)/2}$ as before, and again belong to different systems. Hence

(12) *The groups $G_\pm(n)\,(I_p)$ of the quadrics $Q_\pm(n)$ each contain an invariant subgroup of index two, generated by an even number of I_p's, which leaves each of the two systems of $S_{(n-1)/2}$'s (real or imaginary) on $Q_\pm(n)$ unaltered.*

This enables us to give the complete constitution of the group of the quadric of Dickson's type B:

(13) *The quadric,*

$$B = \nu y_0^2 + y_1^2 + \cdots + y_n^2 \ (n \text{ odd}),$$

[a $Q_+(n)$ when $n \equiv 1$ mod. 4 in G. F. (II), *otherwise a $Q_-(n)$] has a collineation group of order $N_+(n)$ or $N_-(n)$, as the case may be [cf. (3)]. This has an invariant subgroup $G(I_p)$ of index two, generated by elements I_p, whose elements transform points p_o, p_i into points p_o, p_i respectively. This subgroup has an invariant subgroup of index two, generated by an even number of elements I_p, whose elements transform each of the two systems of $S_{(n-1)/2}$'s on the quadric into itself. Thus the group of B has factors of composition,*

$$2,\ 2,\ N_\pm(n)/4 = SO(n+1, p^m) \qquad [cf.\ ^1,\ \text{p. } 191].$$

There remains only Dickson's type A, n odd, which is a $Q_-(n)$ when $n \equiv 1$ mod. 4 in G. F. (II) but which otherwise is a $Q_+(n)$. For these cases we see in (9) that ρ_2 is odd in G. F. (I) and even in G. F. (II).

Though the table (9) indicates invariant subgroups of the groups $G_\pm(I_p)$ in certain cases, it does not completely describe either the type A or the type B. For this purpose we seek the number of conjugate o-pairs and i-pairs under I_{p_o} and under I_{p_i}. Consider then I_{p_o} for the quadric Q. On p_o there are ρ_1 tangents, ρ_2 secants, and ρ_0 skew-lines of Q. On a tangent through p_0 the point p_o is fixed and the contact on Q is fixed. The $s-1$ other points are all o-points [cf. **1** (13)] and yield $(s-1)/2$ o-pairs. On a secant through p_o

the fourth harmonic of p_o as to the two points on Q is a point p'_o in G. F. (I) and a point p_i in G. F. (II) *cf.* [1 (14)]. Hence [cf. 1 (13)] the secant contains $(s-5)/4$ o-pairs and $(s-1)/4$ i-pairs in G. F. (I), and $(s-3)/4$ o-pairs and $(S-3)/4$ i-pairs in G. F. (II). On a skew line through p_o the fourth harmonic of p_o as to the two imaginary points on Q is a point p_i in G. F. (I), and a point p_o in G. F. (II). Hence the skew line contains $(s-1)/4$ o-pairs and $(s-1)/4$ i-pairs in G. F. (I), and $(s-3)/4$ o-pairs and $(s+1)/4$ i-pairs in G. F. (II). Using a similar argument for I_{p_i}, we have the following result:

(14) *The number of pairs of conjugate o-points, and of conjugate i-points of the involution* I_{p_o}, *and the involution* I_{p_i}, *for a quadric* Q *is given by the table:*

$$
\begin{array}{lllr}
I_{p_o}: & \text{G. F. (I)} & : \rho_1(s-1)/2+\rho_2(s-5)/4+\rho_0(s-1)/4 & o\text{-pairs}, \\
 & & \rho_2(s-1)/4+\rho_0(s-1)/4 & i\text{-pairs}, \\
 & \text{G. F. (II)} & : \rho_1(s-1)/2+\rho_2(s-3)/4+\rho_0(s-3)/4 & o\text{-pairs}, \\
 & & \rho_2(s-3)/4+\rho_0(s+1)/4 & i\text{-pairs}, \\
I_{p_i}: & \text{G. F. (I)} & : \rho_2(s-1)/4+\rho_0(s-1)/4 & o\text{-pairs}, \\
 & & \rho_1(s-1)/2+\rho_2(s-5)/4+\rho_0(s-1)/4 & i\text{-pairs}, \\
 & \text{G. F. (II)} & : \rho_2(s-3)/4+\rho_0(s+1)/4 & o\text{-pairs}, \\
 & & \rho_1(s-1)/2+\rho_2(s-3)/4+\rho_0(s-3)/4 & i\text{-pairs},
\end{array}
$$

where the numbers ρ_0, ρ_1, ρ_2 *refer to the point* p_o *or* p_i *in* I_p, *and for given* Q *are obtained from* **1** (12).

We are interested only in the parity of the numbers of the table (14). For $Q_\pm(n)$, $\rho_1 = P_{n-2} = s^{n-2} + \cdots + s + 1$. Since s and n are odd, ρ_1 is even. Since $(s-1)/2$ is integral, $\rho_1(s-1)/2$ is even and may be dropped in (14) without affecting parity. By adding and subtracting ρ_2 in the first, fourth, sixth, and seventh of the above eight numbers the factor $\rho_2+\rho_0 = s^{n-1}$ appears in all. This being odd, it may be replaced by unity. The ρ_2 still remaining may be replaced by ± 1 in G. F. (I), and may be dropped in G. F. (II). Thus the parity of the eight numbers in (14) is that of

$$
(15)\quad
\begin{array}{lll@{\qquad}lll}
I_{p_o}: & \text{G. F. (I)} & : (s-1)/4+1 & I_{p_i}: & \text{G. F. (I)} & : (s-1)/4 \\
 & & (s-1)/4 & & & (s-1)/4+1 \\
 & \text{G. F. (II)} & : (s-3)/4 & & \text{G. F. (II)} & : (s-3)/4+1 \\
 & & (s-3)/4+1 & & & (s-3)/4.
\end{array}
$$

We observe first that if the number of o-pairs is odd, the number of i-pairs is even and vice-versa. We observe also that the parity of any one of the

four numbers for I_{p_o} is opposite to that of the corresponding number for I_{p_i}. Hence

(16) *The quadric,*

$$A = y_0^2 + y_1^2 + \cdots + y_n^2 \ (n \text{ odd}),$$

[*a* $Q_-(n)$ *when* $n \equiv 1$ *mod.* 4 *in* G. F. (II), *otherwise a* $Q_+(n)$] *has a collineation group of order* $N_+(n)$ *or* $N_-(n)$ *as the case may be* [*cf.* (3)]. *This has an invariant subgroup* $G(I_p)$ *of index two, generated by elements* I_p, *whose elements transform points* p_o, p_i *into points* p_o, p_i *respectively. This subgroup has an invariant subgroup of index two, generated by an even number of elements* I_p, *whose elements transform each of the two systems of* $S_{(n-1)/2}$'s *on the quadric into itself. This second invariant subgroup has an invariant subgroup of index two, generated by pairs of involutions* I_{p_o}, *or by pairs* I_{p_i}, *whose elements permute both the outside points and the inside points of the quadric evenly. Thus the group of* A *has factors of composition,* 2, 2, 2,

$$N_\pm(n)/8 = FO(n+1, p^m) \qquad [cf.\ ^1, \text{p. } 191].$$

3. The groups defined by a quadric Q and a linear space L. The space L being the polar of a point p, we shall have to distinguish the three cases in which p is a q-, an o-, or an i-point, and correspondingly L is an L_q, L_o, or L_i space. The order of the collineation group of Q, L is the order of the group of Q divided by the number of points of the kind in question, these values being obtained from **2** (3) and **1** (10) respectively. Thus we find that

$$
(1) \qquad
\begin{array}{rll}
\text{I 1:} & [Q(n), L_q] & = (s-1)s^{n-1} \cdot N(n-2), \\
\text{I 2:} & [Q(n), L_o] & = N_+(n-1), \\
\text{I 3:} & [Q(n), L_i] & = N_-(n-1), \\
\text{II 1:} & [Q_+(n), L_q] & = (s-1)s^{n-1} \cdot N_+(n-2), \\
\text{II 2:} & [Q_+(n), L_{o,i}] & = 4N(n-1), \\
\text{III 1:} & [Q_-(n), L_q] & = (s-1)s^{n-1} \cdot N_-(n-2), \\
\text{III 2:} & [Q_-(n), L_{o,i}] & = 4N(n-1).
\end{array}
$$

In the case of the quadrics we shall be concerned primarily with the groups $G(I_p)$ generated by the I_p's. For these the cases II, III above become

$$
(2) \qquad
\begin{array}{ll}
[G_\pm(n)(I_p), L_q] & = (s-1)s^{n-1} \cdot G_\pm(n-2)(I_p), \\
[G_\pm(n)(I_p), L_{o,i}] & = 2N(n-1).
\end{array}
$$

We first examine the section of Q, a quadric of any one of the three types, by one of its tangent spaces L_q, tangent at q. If q is the next to the last

reference point, a_{n-1}, the space L_q is $x_n = 0$. This space cuts Q in a quadric in the S_{n-1}, $x_n = 0$, with a node at a_{n-1}. This quadric is a point section by a_{n-1} of a proper quadric $Q(n-2)$ in an S_{n-2} in the S_{n-1} but not on a_{n-1}. We may take this to be the S_{n-2}, $x_{n-1} = x_n = 0$. Then $Q(n-2) = Q(x_{n-1} = x_n = 0)$ is in the standard form and is of the same type as Q. Let $Q = (\alpha x)^2$ and $Q(n-2) = (\beta x)^2 = (\alpha x)^2[x_{n-1} = x_n = 0]$.

The lines λ on q and in L_q give rise to the points in S_{n-2}. These lines λ are the tangents and generators of Q on q. The generators give rise to the points of $Q(n-2)$, the tangents to the o- or i-points of $Q(n-2)$ according as to whether outside q they contain only o-points or only i-points of Q [*cf.* **1** (13)]. The collineation group of $[Q, L_q]$ permutes these lines λ. Those collineations of the group which leave each line λ unaltered [the identity in the group of $Q(n-2)$] must form an invariant subgroup of $[Q, L_q]$. The existence of the factor $N(n-2)$ in the order $[Q, L_q]$ suggests that this invariant subgroup has the order $(s-1)s^{n-1}$. We consider the form of the elements,

$$x'_i = \Sigma_j \gamma_{ij} x_j \qquad (i, j = 0, \cdots, n), \tag{3}$$

of this invariant subgroup.

Let the coefficients in the last column of (3) be $y_0, y_1, \cdots, y_n$. Since $Q = (\alpha x)^2$ is to be unaltered, the term in x_n^2 in $(\alpha x')^2$ must not appear, i. e., $(\alpha y)^2 = 0$. Since the lines λ on a_{n-1} are to be invariant, only the $(k+1)$-th and n-th coördinates of the $(k+1)$-th reference point can be affected whence in the $(k+1)$-th column only the coefficients $\gamma_{k,k}$, $\gamma_{n-1,k}$ appear $(k = 0, \cdots, n-2)$. Since also the point a_{n-1} is to be invariant, $\gamma_{n-1,n-1}$ is the only non-zero coefficient in the n-th column. Since the point $1, 1, \cdots, 1, 0, 0$ is to go into $1, 1, \cdots, 1, \mu, 0$, $\gamma_{00} = \gamma_{11} = \cdots = \gamma_{n-2,n-2}$. The determinant of (3) is the product of these γ's and $\gamma_{n-1,n-1}\, y_n$, whence $y_n \neq 0$, and we may take $\gamma_{00} = \cdots = \gamma_{n-2,n-2} = 1$. Since $(\alpha x')^2 = \rho(\alpha x)^2$, we find from the leading term, $x_0^2, x_0 x_1$ or $-\nu x_0^2$, that $\rho = 1$, and then, from the term in $x_{n-1}x_n$, that $\gamma_{n-1,n-1} = 1/y_n$. Again from the absolute invariance of Q we find that

$$\gamma_{n-1,0}x_0 + \cdots + \gamma_{n-1}x_{n-2} = -(y\partial/\partial x)\,Q(n-2)/y_n. \tag{4}$$

The point y, on Q but not on $x_n = 0$, is joined to a_{n-1} by one of the s^{n-1} secants of Q on a_{n-1} [cf. **1** (12)]. Hence

(5) *For any of the s^{n-1} choices of y on Q but not on L_q $(q = a_{n-1}, L_q \equiv x_n = 0)$, and for any non-zero factor of proportionality y_n $[(s-1)$ choices$]$, the ele-*

ment (3) *is a collineation of determinant unity which transforms* Q *into itself and leaves every line* λ *on* q *in* L_q *unaltered if the coefficients satisfy the requirements: the matrix of the first* $n-1$ *rows and columns is the unit matrix, all the other* γ_{ij}*'s are zero except that* $\gamma_{n-1,n-1}=1/y_n$*, the last column is* $y_0,\cdots,y_n$*, and the coefficients* $\gamma_{n-1,0},\cdots,\gamma_{n-1,n-2}$ *are as in* (4). *These* $(s-1)s^{n-1}$ *elements form an invariant subgroup* $g_{(s-1)s^{n-1}}$ *of the group* $[Q,L_Q]$.

The multiplication table of this group is easily obtained. The element described in (5) is uniquely defined by y on Q $(y_n\neq 0)$. Let us call it T_y. If T_z is another element defined by z on Q $(z_n\neq 0)$, the product $T_yT_z=T_t$ is defined by t on Q $(t_n\neq 0)$. To obtain this product we need to get only the last column of coefficients. This yields

$$T_yT_z=T_t \tag{6}$$
$$t_0=y_0+y_nz_0,\ t_1=y_1+y_nz_1,\cdots,t_{n-2}=y_{n-2}+y_nz_{n-2},\ t_{n-1}=\kappa,\ t_n=y_nz_n,$$

where the explicit form of κ is not material, it being uniquely determined by the fact that t is on Q.

We see from (6) that if y_n and z_n are 1, t_n is 1, and $T_yT_z=T_zT_y=T_t$. Also, if T_y, T_z are inverse, $y_nz_n=1$, or $z_n=1/y_n$. Hence, if $y_n=1$, $T_z^{-1}T_yT_z$ has a coefficient $\gamma_{n,n}=1/z_n\cdot 1\cdot z_n=1$. Thus

(7) *The elements of* $g_{(s-1)s^{n-1}}$ *in* (5) *for which* $y_n=1$ *form an abelian subgroup* $g_{s^{n-1}}$ *of order* s^{n-1} [*for* s^{n-1} *choices of* $y_0,\cdots,y_{n-2}$] *invariant under* $g_{(s-1)s^{n-1}}$. *This abelian subgroup is of type* $(1,1,\cdots,1)$, *contains only transformations of period* p, *and is a regular group on the* s^{n-1} *secants of* Q *on* q.

For, any line on a_{n-1} not in $x_n=0$, i. e., any secant, can be *represented* by the point where it cuts $x_{n-1}=0$, i. e., by $s_0,\cdots,s_{n-2},0,1$. This point is transformed by the element (5) into another point whose join with a_{n-1} is *represented* by

$$(s_0+y_0)/y_n,\ (s_1+y_1)/y_n,\cdots,(s_{n-2}+y_{n-2})/y_n,\ 0,\ 1. \tag{8}$$

Thus, when $y_n=1$, we have

$$s'_0=s_0+y_0,\cdots,s'_{n-2}=s_{n-2}+y_{n-2}. \tag{9}$$

Hence, given s and s', there is one and only one set of values $y_0,\cdots,y_{n-2},y_n=1$ for which (9) holds.

According to (6), if $y_n=\epsilon$, a primitive root in G. F., T_y is an element whose $(s-1)$-th power is in the above abelian group, whence

(10) *The factor group of $g_{s^{n-1}}$ with respect to $g_{(s-1)s^{n-1}}$ is a cyclic g_{s-1}.*

From this there follows that

(11) *The factors of composition of the collineation group of order $(s-1)s^{n-1}N(n-2) = [Q, L_q]$ in* (1) I 1, II 1, III 1 *are first the factors of composition of Q_{n-2} [cf.* 2 (10), (13), (16)], *second the factors of a cyclic g_{s-1}, and third the factors of the abelian $g_{s^{n-1}}$ of type* $(1, 1, \cdots, 1)$.

It is clear from the form of (5) that

(12) *Each of the $s^{n-1}-1$ elements of $g_{s^{n-1}}$ other than the identity has an S_{n-2} of fixed points, $x_n = 0$ and $(y\partial/\partial x)Q(n-2)(x) = 0$* [cf. (4)]. *Each of these S_{n-2}'s in the S_{n-1}, $x_n = 0$, on a_{n-1} [$P_{n-2} = (s^{n-1}-1)/(s-1)$ in number] arises from $s-1$ transformations due to the factor in $y_0, \cdots, y_{n-2}$.*

If in (8) we set $s'_i = (s_i + y_i)/y_n$ $(i = 0, \cdots, n-2)$, then there is a fixed secant when $y_n \neq 1$, namely: $s_i = y_i/(y_n - 1)$. Hence

(13) *Any element in $g_{(s-1)s^{n-1}}$ not in $g_{s^{n-1}}$ has one and only one fixed secant on $q = a_{n-1}$. The subgroup of order $s-1$ which leaves one secant fixed is generated by pairs of involutions I_p for points p on the secant but not on Q.*

For, if we take $a_{n-1}a_n$ as a typical secant, this subgroup has the form $x'_i = x_i$ $(i = 0, \cdots, n-2)$, $x'_{n-1} = x_{n-1}/y_n$, $x'_n = y_n x_n$. This is the product of the pair of involutions, $x'_{n-1} = x''_n, x'_n = x''_{n-1}$; $x''_{n-1} = y_n x_n, x''_n = x_{n-1}/y_n$, the other variables being unaltered. For $y_n = -1$, this product is an involution of rank $(1, n-2)$, i. e., with a line and an S_{n-2} of fixed points. On the secant this effects the identity so that the g_{s-1} in (13) is, on the secant, the $g_{(s-1)/2}$ formed by products of an even number of reflections in real pairs apolar to a real pair.

The elements of $g_{(s-1)s^{n-1}}$ in $g_{s^{n-1}}$ may be obtained by taking a product of two products of pairs of involutions for one of which we have y_n, and for the other $z_n = 1/y_n$ [*cf.* (6)]. Hence

(14) *The group $g_{(s-1)s^{n-1}}$ is generated by products $I_p I_{p'}$ where p, p' are any two points not on Q but on a secant through q.*

With respect to the problem outlined in the introduction however, we want the group which leaves Q, L absolutely unaltered, and which is generated by involutions I_p for points p *on* the linear space L. Thus, our linear space being $x_n = 0$, this requirement restricts the transformations (5) for which

$x'_n = y_n x_n$ to those for which $y_n = 1$, i. e., to the subgroup $g_{s^{n-1}}$ of $g_{(s-1)s^{n-1}}$. On $q = a_{n-1}$ the tangents τ and generators γ of Q are cut by any S_{n-1} not on a_{n-1} in the points respectively, not on, and on, the quadric $Q(n-2)$. The points p on L but not on Q are points p_τ on these tangents. The involutions, $I_{p\tau}, I_{p'\tau}$ (p, p' on the same tangent τ), effect the same permutation on tangents τ and generators γ, when $I_{p\tau}I_{p'\tau}$ leaves each unaltered, and also leaves Q, L unaltered, and thus is in $g_{s^{n-1}}$. Conversely $g_{s^{n-1}}$ is generated by such pairs. To prove this let $z = z_0, \cdots, z_{n-2}, 0, 0$. If then $(\beta z)^2 \neq 0$, the point $p_\tau = z + \mu a_{n-1}$ $(\mu \neq 0)$ is a point on the tangent τ. The involution $I_{p\tau}$ is [cf. 2 (4)]

$$x' = x - (z + \mu a_{n-1})[2(\beta z)(\beta x) + \mu x_n]/(\beta z)^2.$$

If $I_{p'\tau}$ is another point on the same tangent determined by $\mu' \neq \mu$, then

$$\begin{aligned} I_{p\tau}I_{p'\tau} : x' = x &+ z \cdot (\mu - \mu')x_n/(\beta z)^2 \\ &+ a_{n-1} \cdot (\mu' - \mu)[2(\beta z)(\beta x) + (\mu - \mu')x_n]/(\beta z)^2. \end{aligned}$$

If we compare this with (5) we find that

$$y_0, \cdots, y_n = \begin{matrix} -(\mu - \mu')^2/(\beta z)^2, 1, \\ (\mu - \mu')z_0/(\beta z)^2, \cdots, (\mu - \mu')z_{n-2}/(\beta z)^2, \end{matrix}$$

and thus $I_{p\tau}I_{p'\tau}$ is in $g_{s^{n-1}}$ since $y_n = 1$. We do not however find in this way all the elements of $g_{s^{n-1}}$ due to the restriction, $(\beta z)^2 \neq 0$. If then we choose z, z' not on $Q(n-2)$ such that $z + z' = t$ is on $Q(n-2)$, the element of $g_{s^{n-1}}$ corresponding to $y = t$ is $I_{p\tau}I_{p'\tau}I_{q\tau'}I_{q'\tau'}$ where τ' is the tangent determined by z'. Hence

(15) *The group $g_{s^{n-1}}$ is generated by products of pairs of involutions I_p for pairs of points p on tangents τ of Q in L_q.*

The involution $I_{p\tau}$ effects the same permutation on lines λ on a_{n-1} in L as the involution I_p for $Q(n-2)$ effects on the points of S_{n-2}. Hence these involutions generate the $G(n-2)(I_p)$ on the lines λ and in S_n the $G(I_{p\tau})$ has the invariant $g_{s^{n-1}}$ with factor group $G(n-2)(I_p)$. Hence

(16) *The collineation group, $[G(I_p), L_q]$, generated by involutions I_p for points p on L_q, and thus leaving Q, L_q unaltered, has factors of composition which are the factors of $G(n-2)(I_p)$ [cf. 2 (10), (13), (16)], and the factors of $g_{s^{n-1}}$.*

We now consider the cases (1) $I(2)I(3)$. Given $Q(n)$ and an outside

point o, the polar space L_o cuts $Q(n)$ in a $Q_+(n-1)$. The order of the entire collineation group of $Q(n)$, L_o is $N_+(n-1)$. This we write as $2 \cdot N_+(n-1)/2$, since we are interested only in collineations which leave $Q(n)$, L_o, and therefore the section $Q_+(n-1)$, absolutely unaltered. The latter group is generated by involutions I_p for points p on L_o. But the involutions attached to n linearly independent points in L_o generate the involution I_o which necessarily is invariant under the entire group. Similar remarks apply to an inside point i and the section $Q_-(n-1)$. Hence

(17) *The collineation group,* $[G(n)(I_p), L_o]$ $\{[G(n)(I_p), L_i]\}$ *generated by involutions* I_p *for points* p *on* $L_o\{L_i\}$ *has for factors of composition those of the group* $G_+(n-1)(I_p)$ $\{G_-(n-1)(I_p)\}$, *and a factor corresponding to the invariant* g_2 *generated by* $I_o\{I_i\}$.

There remain finally the cases II 2, III 2. We pass immediately to the $G(I_p)$ and, in connection with (2), state that:

(18) *The collineation group* $[G_\pm(n)(I_p), L_{o,i}]$ *generated by involutions* I_p *for points* p *on* L_o *or* L_i, *has for factors of composition those of the group* $G(n-1)(I_p)$ *and a factor* 2 *corresponding to the invariant* g_2 *generated by* I_o *or* I_i.

REFERENCES.

[1] L. E. Dickson, *Linear Groups*, Leipzig (Teubner), 1901.

[2] C. Segre, *Mem. Acc. Torino* (1884).

[3] E. Bertini, *Introduzione alla Geometria Projettiva degli Iperspazi*, Pisa (1907), pp. 133-135.

[4] A. B. Coble, "Algebraic geometry and theta functions," *Colloquium Publications of the American Mathematical Society*, vol. 10 (1929).

[5] A. B. Coble, "Theta modular groups determined by point sets," *American Journal of Mathematics*, vol. 40 (1918), pp. 317-340.

ON THE DISTRIBUTION OF THE VALUES OF THE RIEMANN ZETA FUNCTION.

By H. Bohr and B. Jessen.

Introduction. The object of this note is to fill in a gap in the description of the distribution of the values of the Riemann zeta function $\zeta(s) = \zeta(\sigma + it)$, or rather the function $\log \zeta(s)$, in the half-plane $\sigma > 1$.[1] From the Euler product we have for this function the expression

$$-\log \zeta(s) = \sum_{n=1}^{\infty} \log (1 - p_n^{-s}) = \sum_{n=1}^{\infty} \log (1 - p_n^{-\sigma} e^{-it \log p_n}),$$

where p_n denotes the prime numbers $2, 3, 5, \cdots$.[2]

For a fixed $\sigma > 1$ we consider in the complex w-plane the closure $M(\sigma)$ of the set of values $-\log \zeta(\sigma + it)$, $-\infty < t < +\infty$. It is known that, on account of the linear independence of the numbers $\log p_n$, this set $M(\sigma)$ is identical with the range of values of the function

$$F(\theta_1, \theta_2, \cdots) = \sum_{n=1}^{\infty} \log (1 - p_n^{-\sigma} e^{i\theta_n}),$$

where the real variables $\theta_1, \theta_2, \cdots$ are independent of each other and θ_n describes the interval $0 \leqq \theta_n < 2\pi$. Thus, if for an arbitrary r in $0 < r < 1$ we denote by $S(r)$ the curve

(1) $$w = \log (1 + x), \quad |x| = r,$$

we have

(2) $$M(\sigma) = \sum_{n=1}^{\infty} S(p_n^{-\sigma}),$$

where the sum of sets is to be taken in the vectorial sense, that is, the sum denotes the set of all points $w = \sum_{n=1}^{\infty} w_n$, where w_n belongs to $S(p_n^{-\sigma})$.

---

[1] The results concerning the distribution of the values of the function $\log \zeta(s)$ in the half-plane $\sigma > 1$ which we shall use and which are restated in the text are given in H. Bohr [2]. For a more detailed study of the distribution of the values involving also problems of probability and dealing with the half-plane $\sigma > \frac{1}{2}$, we refer to H. Bohr and B. Jessen [5-6], and, particularly, to the comprehensive treatment in B. Jessen and A. Wintner [9]. The results concerning the function $\zeta'(s)/\zeta(s)$ which we mention are given in H. Bohr [1] and C. Burrau [7].

[2] We consider $-\log \zeta(s)$ instead of as usual $\log \zeta(s)$ itself in order to avoid the minus sign in several other places. We notice that all occurring infinite series are absolutely convergent.

From this representation of $M(\sigma)$ and the simple fact, to which we return below, that each curve $S(r)$ is convex, has been obtained the following simple result concerning the shape of the set $M(\sigma)$: that $M(\sigma)$ is for each $\sigma > 1$ either a closed domain bounded by a single convex curve $A(\sigma)$ or a closed ring-shaped domain, bounded by two convex curves $A(\sigma)$ and $B(\sigma)$, where $B(\sigma)$ lies inside $A(\sigma)$. Furthermore, it was shown by rough estimations that for all σ sufficiently near to 1 the first case occurs, while the second case occurs for all sufficiently large σ.[3]

So far the situation is quite similar to that obtained for the derivative $\zeta'(s)/\zeta(s)$ of the function $\log \zeta(s)$, only in this latter case the situation is simplified by the fact that the convex curves to be added turn out to be circles. Their sum is therefore either the closed surface of a circle or a closed concentric circular ring. In this case it was possible by simple computations to decide for which σ each of the two cases occurred, the result being the existence of a constant $D > 1$, such that for $\sigma \leqq D$ the case of the circle, for $\sigma > D$ the case of the circular ring takes place. A numerical calculation gave the approximate value $D = 2.576076$.

The object of the present note is to prove that a quite analogous situation holds for the function $\log \zeta(s)$ itself.

THEOREM. *Denoting by $M(\sigma)$ for $\sigma > 1$ the closure of the set of values $-\log \zeta(\sigma + it)$, $-\infty < t < +\infty$, there exists a constant $C > 1$, so that for each $\sigma \leqq C$ the set $M(\sigma)$ is a closed domain bounded by a single convex curve $A(\sigma)$, while for each $\sigma > C$ the set $M(\sigma)$ is a closed ring-shaped domain bounded by two convex curves $A(\sigma)$ and $B(\sigma)$, where $B(\sigma)$ lies inside $A(\sigma)$.*

C is characterized as the only root in $\sigma > 1$ of the equation

$$\operatorname{arc\,sin} 2^{-\sigma} = \sum_{n=2}^{\infty} \operatorname{arc\,sin} p_n^{-\sigma}.$$

We have the approximate value $C = 1.764$, correct to two decimal places.

Some details regarding the shape of the curve $B(\sigma)$ for $\sigma > C$ are contained in a theorem at the end of this note.

A geometrical criterion that the set be ring-shaped. In our investigations of the set $M(\sigma)$ as given by the formula (2) we shall make no use of any special properties of the prime numbers p_n except that $1 < p_1 < p_2 < \cdots$

[3] For the problem of the addition of convex curves see H. Bohr [3], H. Bohr and B. Jessen [4] and E. K. Haviland [8]. A short account is to be found in B. Jessen and A. Wintner [9].

and that the series $\sum_{n=1}^{\infty} p_n^{-\sigma}$ converges for $\sigma > 1$ and diverges for $\sigma = 1$. It will therefore be more natural to consider the general case, where the sets $M(\sigma)$ for $\sigma > 1$ are defined by

$$(3) \qquad M(\sigma) = \sum_{n=1}^{\infty} S(e^{-\lambda_n \sigma}),$$

where $0 < \lambda_1 < \lambda_2 < \cdots$ is any sequence such that the series $\sum_{n=1}^{\infty} e^{-\lambda_n \sigma}$ converges for $\sigma > 1$ and diverges for $\sigma = 1$.[4]

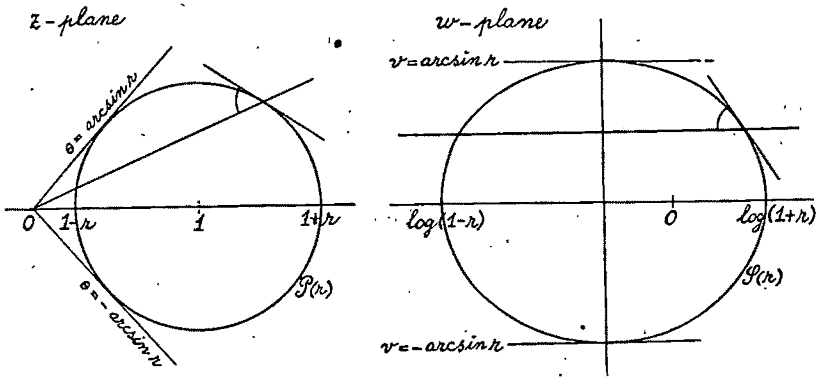

FIGURE 1.

We begin with some simple remarks concerning the curve $S(r)$, $0 < r < 1$, given by the representation (1). When x describes the circle $|x| = r$, the point $z = 1 + x$ describes the circle $P(r)$ in the z-plane with mid-point 1 and radius r (see Figure 1). The curve $S(r)$ is the image of this circle $P(r)$ by the conformal representation $w = \log z$. Writing $z = \rho e^{i\theta}$, $-\pi/2 < \theta < \pi/2$, and $w = u + iv$, we have

$$u = \log \rho \quad \text{and} \quad v = \theta.$$

The ranges of ρ and θ, when z describes $P(r)$, are

$$1 - r \leqq \rho \leqq 1 + r \quad \text{and} \quad -\arcsin r \leqq \theta \leqq \arcsin r.$$

[4] Evidently the constant C occurring in the theorem will depend on the sequence $\lambda_1, \lambda_2, \cdots$. We notice that in the special case, where the λ_n are linearly independent, the set $M(\sigma)$ given by (3) is the closure of the set of values $f(\sigma + it)$, $-\infty < t < +\infty$, where $f(s)$ is the analytic function

$$f(s) = \sum_{n=1}^{\infty} \log (1 - e^{-\lambda_n s}).$$

Thus our theorem holds for any function of this kind.

The extreme values of ρ are taken at the points of $P(r)$ lying on the real axis and correspond to the value $\theta = 0$; the extreme values of θ are taken at the touching points of the tangents from the origin to $P(r)$ and correspond to the value $\rho = (1 - r^2)^{\frac{1}{2}}$. To any value of ρ except the extreme ones correspond two values of θ, which are numerically equal but with opposite signs, while to each value of θ except the extreme ones correspond two values of ρ, whose product is $1 - r^2$. Consequently, the curve $S(r)$ is symmetrical with respect to the lines $v = 0$ and $u = \frac{1}{2} \log (1 - r^2)$ and is cut by each line $u = u_0$, $\log (1 - r) < u_0 < \log (1 + r)$, and by each line $v = v_0$, $-\arcsin r < v_0 < \arcsin r$, in exactly two points. Furthermore, the curve $S(r)$ is convex; in fact, the angle between the tangent of $S(r)$ at the point (u_0, v_0) and the line $v = v_0$ is, in the virtue of the conformity, equal to the angle between the tangent of $P(r)$ at the corresponding point (ρ_0, θ_0) and the line $\theta = \theta_0$, and this latter angle varies monotonously when the point describes $P(r)$.[5]

By $S_0(r)$ we shall denote the curve obtained from $S(r)$ by the translation $-\frac{1}{2} \log (1 - r^2)$, having the symmetry axes $v = 0$ and $u = 0$. It is obvious that, if $0 < r'' < r' < 1$, the image $S(r'')$ of the circle $P(r'')$ lies inside the image $S(r')$ of the circle $P(r')$. In virtue of the symmetry and convexity of the two curves this implies that $S_0(r'')$ must lie inside $S_0(r')$.

The symmetry of the single curve $S(r)$ immediately involves a similar symmetry of the set $M(\sigma)$ defined by (3), the axes of symmetry being $v = 0$ and $u = \sum_{n=1}^{\infty} \frac{1}{2} \log (1 - e^{-2\lambda_n\sigma})$. By $M_0(\sigma)$ we shall denote the set obtained from $M(\sigma)$ by the translation $-\sum_{n=1}^{\infty} \frac{1}{2} \log (1 - e^{-2\lambda_n\sigma})$, having the symmetry axes $v = 0$ and $u = 0$. Obviously

$$M_0(\sigma) = \sum_{n=1}^{\infty} S_0(e^{-\lambda_n\sigma}).$$

We shall now deduce a simple criterion enabling us to decide whether for a given σ the set $M_0(\sigma)$ and hence also the set $M(\sigma)$ is ring-shaped or convex. For this purpose we write $M_0(\sigma)$ in the form

$$M_0(\sigma) = S_0(e^{-\lambda_1\sigma}) + N_0(\sigma), \quad \text{where} \quad N_0(\sigma) = \sum_{n=2}^{\infty} S_0(e^{-\lambda_n\sigma}).$$

---

[5] The two half axes of $S(r)$ lying on the two axes of symmetry $v = 0$ and $u = \frac{1}{2} \log (1 - r^2)$ are

$$a(r) = \tfrac{1}{2} [\log (1 + r) - \log (1 - r)] \quad \text{and} \quad b(r) = \arcsin r$$

respectively. For the orientation of the reader, we notice without proof that the ratio $a(r)/b(r)$ is an increasing function of r in $0 < r < 1$, so that the curve $S(r)$, which for small values of r is approximately a circle, becomes more and more oblong as r increases.

As the set $M_0(\sigma)$ has the origin as center of symmetry, it is clear that it is convex or ring-shaped according as it contains or does not contain the origin, that is, according as there exist or do not exist points w_1 and w_2 of $S_0(e^{-\lambda_1\sigma})$ and $N_0(\sigma)$ respectively, such that $w_1 + w_2 = 0$, which by the symmetry of either set with respect to the origin is the case according as $S_0(e^{-\lambda_1\sigma})$ and $N_0(\sigma)$ have or have not points in common.

Now as a sum of convex curves the set $N_0(\sigma)$ is itself either a closed domain bounded by a simple convex curve $C_0(\sigma)$ or a closed ring-shaped domain bounded by two convex curves $C_0(\sigma)$ and $D_0(\sigma)$, where $D_0(\sigma)$ lies inside $C_0(\sigma)$. From the general considerations regarding the addition of convex curves and the fact that all the curves $S_0(e^{-\lambda_n\sigma})$, $n \geqq 3$, surround the origin and are contained in $S_0(e^{-\lambda_2\sigma})$, it follows immediately that the interior boundary $D_0(\sigma)$ of $N_0(\sigma)$, if it exists, must lie inside $S_0(e^{-\lambda_2\sigma})$ and hence still more inside $S_0(e^{-\lambda_1\sigma})$. Thus in the determination, whether $S_0(e^{-\lambda_1\sigma})$ and $N_0(\sigma)$ have or have not points in common, it will make no difference if, in the case where $D_0(\sigma)$ exists, we add to $N_0(\sigma)$ the interior of $D_0(\sigma)$, so that in all cases the problem is only to decide whether $S_0(e^{-\lambda_1\sigma})$ and the closed convex domain bounded by the curve $C_0(\sigma)$ have or have not points in common. Hence we have the following criterion:

The set $M_0(\sigma)$ is ring-shaped or convex according as the convex curve $C_0(\sigma)$ lies or does not lie inside $S_0(e^{-\lambda_1\sigma})$.

An analytical formulation of the criterion. That a convex curve lies inside another convex curve may be expressed analytically in various ways. For the present purpose, where one of the two curves to be considered is given as the exterior boundary for a sum of convex curves, the obvious procedure is to use the supporting functions (Stützfunktion) of the two curves, since the supporting function of the exterior boundary of a sum of convex curves is immediately expressed as the sum of the supporting functions for the single curves to be added.

For each r in $0 < r < 1$ we denote by $H_0(r; \alpha)$, where α is an angular variable, the supporting function of the convex curve $S_0(r)$, defined as the maximum of $u \cos \alpha + v \sin \alpha$ when (u, v) describes $S_0(r)$. The explicit expression for $H_0(r; \alpha)$ is somewhat complicated, but will not be needed; we shall use only the special value $H_0(r; \pi/2) = \arcsin r$. Denoting by $K_0(\sigma; \alpha)$ the supporting function of the curve $C_0(\sigma)$, we have

$$K_0(\sigma; \alpha) = \sum_{n=2}^{\infty} H_0(e^{-\lambda_n\sigma}; \alpha).$$

Now if S' and S'' are two convex curves and $H'(\alpha)$ and $H''(\alpha)$ their supporting functions, the necessary and sufficient condition that S'' lie inside S' is that $H'(\alpha) > H''(\alpha)$ for all α. We may therefore state the above criterion in the following form:

The set $M(\sigma)$ is ring-shaped or convex according as the inequality

$$H_0(e^{-\lambda_1\sigma}; \alpha) > \sum_{n=2}^{\infty} H_0(e^{-\lambda_n\sigma}; \alpha) \tag{4}$$

holds or does not hold for all α. For reasons of symmetry it is evidently sufficient to consider values $0 \leqq \alpha \leqq \pi/2$.

Proof of the theorem. By virtue of the last criterion our theorem will follow from the following two propositions:

(i) *For any fixed $\sigma > 1$ the inequality* (4) *will hold for all α, if it holds for $\alpha = \pi/2$.*

(ii) *There exists a constant $C > 1$ such that the inequality*

$$H_0(e^{-\lambda_1\sigma}; \pi/2) > \sum_{n=2}^{\infty} H_0(e^{-\lambda_n\sigma}; \pi/2)$$

is or is not satisfied, according as $\sigma > C$ or $\sigma \leqq C$.

Proof of proposition (i). For a fixed $\sigma > 1$ we put $e^{-\lambda_n\sigma} = r_n$, so that $r_1 > r_2 > \cdots$. The proposition to be proved is then that the inequality

$$H_0(r_1; \alpha) > \sum_{n=2}^{\infty} H_0(r_n; \alpha)$$

holds for $0 \leqq \alpha < \pi/2$, if it holds for $\alpha = \pi/2$. Obviously this will be proved if we prove that for $0 < r'' < r' < 1$ the inequality

$$\frac{H_0(r''; \alpha)}{H_0(r''; \pi/2)} < \frac{H_0(r'; \alpha)}{H_0(r'; \pi/2)},$$

that is, the inequality

$$\frac{H_0(r''; \alpha)}{\arcsin r''} < \frac{H_0(r'; \alpha)}{\arcsin r'},$$

holds for $0 \leqq \alpha < \pi/2$.

Denoting for an arbitrary r in $0 < r < 1$ by $S^*_0(r)$ the convex curve similar to $S_0(r)$ with respect to the origin in the ratio $1/(\arcsin r)$, the supporting function $H^*_0(r; \alpha)$ of $S^*_0(r)$ is

$$H^*_0(r; \alpha) = \frac{H_0(r; \alpha)}{\arcsin r},$$

so that the assertion to be proved is that for $0 < r'' < r' < 1$ the inequality

$$H^*_0(r''; \alpha) < H^*_0(r'; \alpha)$$

holds for $0 \leqq \alpha < \pi/2$. This inequality simply expresses that the curve $S^*_0(r'')$ lies inside $S^*_0(r')$ except for the two points $(0, 1)$ and $(0, -1)$, which lie on all the curves $S^*_0(r)$.[6]

For an arbitrary r in $0 < r < 1$ we represent the part of the curve $S^*_0(r)$ lying in $u \geqq 0$, $v \geqq 0$ by an equation $u = f(r; v)$, $0 \leqq v \leqq 1$. Denoting for a fixed t in $0 < t < 1$ by $\alpha(r; t)$ the angle between the line $v = t$ and the normal of the curve $S^*_0(r)$ at the point $(f(r; t), t)$, we have $\tan \alpha(r; t) = -f'_v(r; t)$, so that

$$f(r; v) = \int_v^1 \tan \alpha(r; t)\, dt$$

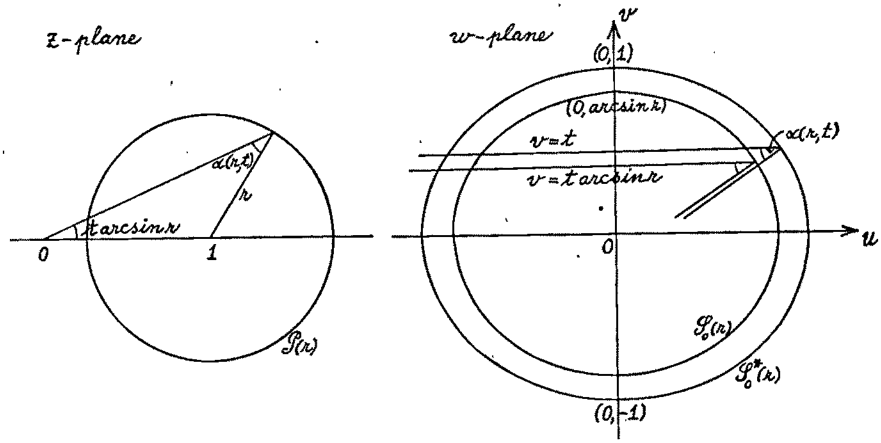

FIGURE 2.

for $0 \leqq v < 1$. Our assertion being that for $0 < r'' < r' < 1$ we have $f(r''; v) < f(r'; v)$ in $0 \leqq v < 1$, it is therefore sufficient to prove that for any fixed t in $0 < t < 1$ the angle $\alpha(r; t)$ is an increasing function of r in $0 < r < 1$.

In order to calculate the angle $\alpha(r; t)$ we notice that $\alpha(r; t)$ is also the angle between the line $v = t \arcsin r$ and the normal of the curve $S_0(r)$ or $S(r)$ at one of the points lying on this line, and hence by the conformal representation equal to the angle between the line $\theta = t \arcsin r$ and the normal of the circle $P(r)$ at one of the points lying on this latter line (see Figure 2). Hence we have the relation

---

[6] This contains the statement in the preceding footnote, that the curve $S(r)$ becomes more and more oblong as r increases, and shows that this even happens in a very regular way.

$$\frac{\sin \alpha(r;t)}{1} = \frac{\sin (t \arcsin r)}{r}.$$

Introducing $\arcsin r = y$ as new variable instead of r, we find

$$\sin \alpha(r;t) = \frac{\sin ty}{\sin y}.$$

Thus in order to prove that for any fixed t in $0 < t < 1$ the angle $\alpha(r;t)$ is an increasing function of r in $0 < r < 1$, we have to prove only that

$$g(y) = \frac{\sin ty}{\sin y}$$

is an increasing function of y in $0 < y < \pi/2$, which is clear since

$$g'(y) = \frac{ty \cos y \cos ty}{\sin^2 y}\left(\frac{\tan y}{y} - \frac{\tan ty}{ty}\right) > 0,$$

the function $\tan x/x$ being increasing in $0 < x < \pi/2$.

Proof of proposition (ii). Proposition (ii) states the existence of a constant $C > 1$, such that the function

$$f(\sigma) = \frac{\sum\limits_{n=2}^{\infty} H_0(e^{-\lambda_n \sigma}; \pi/2)}{H_0(e^{-\lambda_1 \sigma}; \pi/2)} = \frac{\sum\limits_{n=2}^{\infty} \arcsin e^{-\lambda_n \sigma}}{\arcsin e^{-\lambda_1 \sigma}}$$

is $\geqq 1$ for $\sigma \leqq C$ and < 1 for $\sigma > C$. Since $f(\sigma) \to \infty$ as $\sigma \to 1$ and $f(\sigma) \to 0$ as $\sigma \to \infty$ on account of the assumptions concerning the λ_n, it is sufficient to prove that the function $f(\sigma)$ is decreasing in $1 < \sigma < \infty$. We shall even prove that each term

$$g(\sigma) = \frac{\arcsin e^{-\lambda_n \sigma}}{\arcsin e^{-\lambda_1 \sigma}}$$

has a negative derivative $g'(\sigma)$ and hence is decreasing in $1 < \sigma < \infty$. The logarithmic derivative $g'(\sigma)/g(\sigma)$ of the function $g(\sigma)$ being the difference between the logarithmic derivatives of the numerator and the denominator, the inequality $g'(\sigma) < 0$ will be proved if we prove that the logarithmic derivative $h'(\sigma)/h(\sigma)$ of the function

$$h(\sigma) = \arcsin e^{-\lambda\sigma}$$

is for each fixed σ in $1 < \sigma < \infty$ a decreasing function of λ in $0 < \lambda < \infty$. Now

$$\frac{h'(\sigma)}{h(\sigma)} = \frac{-\lambda e^{-\lambda\sigma}}{\sqrt{1 - e^{-2\lambda\sigma}} \arcsin e^{-\lambda\sigma}}.$$

For a fixed σ we introduce $\arcsin e^{-\lambda\sigma} = y$ as new variable instead of λ, and have then to prove that the function

$$\phi(y) = \frac{1}{\sigma}\,\frac{\sin y \log \sin y}{\cos y \cdot y} = \frac{1}{\sigma}\,\frac{\tan y \log \sin y}{y} = \frac{1}{\sigma}\,\frac{\psi(y)}{y}$$

is increasing in the interval $0 < y < \pi/2$. Since $\psi(0) = 0$, this will certainly be the case if $\psi(y)$ is convex in $0 < y < \pi/2$, so that it is sufficient to prove that

$$\psi'(y) = 1 + \frac{\log \sin y}{\cos^2 y}$$

is increasing in $0 < y < \pi/2$. Once more changing the variable, putting $\sin y = t$, we have thus only to show that

$$\chi(t) = \frac{\log t}{1 - t^2}$$

is increasing in $0 < t < 1$. But this is clear since

$$\chi'(t) = \frac{1 - t^2 + 2t^2 \log t}{t(1-t^2)^2} = \frac{\xi(t)}{t(1-t^2)^2} > 0$$

in $0 < t < 1$, as $\xi'(t) = 4t \log t < 0$ in $0 < t < 1$ and $\xi(1) = 0$.

Another theorem. It is easily seen that the exterior boundary $A(\sigma)$ of the set $M(\sigma)$ for an arbitrary $\sigma > 1$ contains neither corners nor straight segments. By arguments similar to those applied above, we are able to prove the following theorem regarding the shape of the interior boundary $B(\sigma)$ of the set $M(\sigma)$ for $\sigma > C$:

THEOREM. *There exists a constant $E > C$ such that for each $\sigma < E$ the curve $B(\sigma)$ has exactly two corners lying on the real axis, while for each $\sigma \geqq E$ the curve $B(\sigma)$ has no corners. For no value of σ does $B(\sigma)$ contain straight segments.*

E is characterized as the only root in $\sigma > 1$ of the equation

$$2^{-\sigma} = \sum_{n=2}^{\infty} p_n^{-\sigma}.$$

We have the approximate value $E = 1.778$, correct to two decimal places.

We are indebted to Mr. J. P. Møller for the numerical calculation of the constants C and E.

FYNSHAV, ALS, DENMARK.

REFERENCES.

[1] H. Bohr, "Über die Funktion $\zeta'(s)/\zeta(s)$," *Journal für Mathematik*, vol. 141 (1912), pp. 217-234.

[2] ———, "Sur la fonction $\zeta(s)$ dans le demi-plan $\sigma > 1$," *Comptes Rendus*, vol. 154 (1912), pp. 1078-1081.

[3] ———, "Om Addition af uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab, Forhandlinger*, 1913, pp. 325-366.

[4] H. Bohr and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition af konvekse Kurver," *Danske Videnskabernes Selskab, Skrifter*, (8), vol. 12, no. 3 (1929).

[5-6] ———, "Über die Werteverteilung der Riemannschen Zetafunktion, I-II," *Acta Mathematica*, vol. 54 (1930), pp. 1-35; vol. 58 (1932), pp. 1-55.

[7] C. Burrau, "Numerische Lösung der Gleichung $(2^{-D}\log 2)/(1-2^{-2D}) = \sum_{n=2}^{\infty}[(p_n^{-D}\log p_n)/(1-p_n^{-2D})]$, wo p_n die Reihe der Primzahlen von 3 an durchläuft," *Journal für Mathematik*, vol. 142 (1912), pp. 51-53.

[8] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[9] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

ON A CLASS OF FOURIER TRANSFORMS.

By Aurel Wintner.

The present paper deals with the Fourier analysis of certain analytic functions. § 1 collects the tools to be used. Theorems I and II of § 2 concern the Fourier analysis of a class of meromorphic functions which are reciprocal values of even transcendental entire functions $f(z)$ of genus 0 in z^2. This theorem is suggested by Hausdorff's remarks on the function $1/\mathrm{Cosh}\, t$ (cf. Hausdorff [8]). Theorem VI furnishes the general analytical background to the well known Fourier representation of $|\Gamma(b+it)|^2$, where $b>0$, in the same manner as Theorem I yields the general background of the standard Fourier representation of $1/\mathrm{Cosh}\, t$. Theorem VII is another analogue of Theorem I and concerns the case of an "erhöht" genus 0 in the sense of Pólya (cf. [19], where further references are given). While Theorem I leads, as shown by Theorem III, to a strange consequence of Riemann's hypothesis, Theorem IV is independent of this hypothesis. Theorem V deals with an interesting transcendant defined by a Bernoulli convolution. The elementary Theorem X and the remarks which follow it treat distributions derived by projection from equidistributions which belong to the interior or to the boundary of an n-dimensional sphere. Theorems VIII and IX contain new information about a class of transcendants introduced by Cauchy [4] which play an important rôle in the investigations of Lévy and Pólya on the foundations of the analytic theory of probability; the same transcendants also occur in the Hardy-Littlewood treatment of Waring's problem (for references cf. Pólya [18]). It is shown by Theorems XX and XXI that Theorems VIII and IX may be extended from the case of trigonometrical integrals to the case of integrals containing Bessel functions. As an application of Theorem VIII, it is shown in § 3 that the three standard postulates in the theory of error distribution, which go back to Gauss [6], are not independent of each other, one of them being implied by the two others. As shown by Theorem XIX, a corresponding result holds in the multidimensional case also. The other results proven at the end of § 4 are known in the case $n=1$, where n is the dimension number (cf. Lévy [14]; Pólya [17], Wintner [23]). The difference between the cases $n=1$ and $n>1$ is about the same as that between an ordinary and a partial differential equation. It turns out, however, that the arbitrary function contained in the solution of the problem in the case $n>1$ must be a constant due to the "boundary condition" of a finite

dispersion. This result might have some physical interest in view of the Maxwell assumption on velocity distribution at random ($n = 3$).

INDEX.

1. *The Fourier-Stieltjes transform.* Let $\sigma = \sigma(x)$, $-\infty < x < +\infty$, be a distribution function, i. e., a monotone function for which $\sigma(-\infty) = 0$ and $\sigma(+\infty) = 1$. It may be assumed that

$$\sigma(x) = \tfrac{1}{2}\{\sigma(x+0) + \sigma(x-0)\}$$

holds also when x belongs to the sequence of discontinuity points (if any). If c is a positive constant, $\sigma(cx)$ also is a distribution function; it will be termed similar to $\sigma(x)$. The function $1 - \sigma(-x)$ also is a distribution function and may be called the conjugate of $\sigma(x)$. The conjugate of the conjugate of σ is σ itself. σ will be said to be symmetric if it is identical with its conjugate, i. e., if

$$\sigma(x) + \sigma(-x) \equiv 1.$$

If σ is symmetric and if the curve $\sigma = \sigma(x)$ is concave (from below) in the open interval $0 < x < +\infty$, hence convex in the open interval $-\infty < x < 0$, then σ will be termed a convex distribution function. It is well known that if a function is convex and bounded in an open interval, it is absolutely continuous in this interval. Thus a convex distribution function $\sigma(x)$ is absolutely continuous in every interval not containing $x = 0$. That a convex distribution function may be discontinuous at $x = 0$, is shown by the example $\chi(x) = \frac{1}{2}(1 + \operatorname{sg} x)$, where $\operatorname{sg} x = -1$, 0 or 1 according as $x < 0$, $= 0$ or > 0. This distribution function χ plays the rôle of the unit in what may be called the algebra of distribution functions.

A sequence $\{\sigma_n\}$ of distribution functions is said to be convergent if there exists a distribution function σ such that $\sigma_n(x) \to \sigma(x)$ at every continuity point x of σ. This is what will be meant by writing $\sigma_n \to \sigma$. Thus $\sigma_n \to \sigma$ and $\sigma_n \to \rho$ imply $\sigma = \rho$. It is clear that if every σ_n is symmetric, then so is σ and that if every σ_n is convex in the sense defined above, then so is σ.

The spectrum of a distribution function σ is defined as the set of those

points $x = x_0$ for which $\sigma(x') < \sigma(x'')$ whenever $x' < x_0 < x''$. This terminology is in accordance with the usual physical terminology concerning the frequency distribution determined by σ, i. e., with the Wirtinger-Hilbert terminology in the theory of linear differential and integral equations and is, therefore, at variance with a terminology recently proposed by Wiener [21], p. 163. The discontinuity points of σ, if any, clearly belong to the spectrum. If σ is analytic, then the spectrum is the whole x-axis. If there exists a finite or infinite sequence $\{x_n\}$ of distinct points such that

$$\sum_n [\sigma(x_n + 0) - \sigma(x_n - 0)] = \int_{-\infty}^{+\infty} d\sigma(x), \text{ i. e., } = 1,$$

then σ will be termed purely discontinuous. The spectrum may be the whole x-axis in this case also.

If two independent random variables ξ_1, ξ_2 have the distribution functions σ_1, σ_2, then the distribution function of $\xi_1 + \xi_2$ is denoted by $\sigma_1 * \sigma_2$ and is termed the convolution ("Faltung") of σ_1 and σ_2; it is represented at its continuity points x by the integral

$$\int_{-\infty}^{+\infty} \sigma_1(x - y)\, d\sigma_2(y) \quad .$$

(cf., e. g., Hausdorff [8]). It is easy to see that $\sigma_1 * \sigma_2 = \sigma_2 * \sigma_1$ and $(\sigma_1 * \sigma_2) * \sigma_3 = \sigma_1 * (\sigma_2 * \sigma_3)$. Also, $\sigma * \chi = \sigma$ for any σ, where again $\chi(x) = \frac{1}{2}(1 + \operatorname{sg} x)$. The conjugate of $\sigma_1 * \sigma_2$ is the convolution of the conjugates of σ_1 and of σ_2. It follows that if σ_1 and σ_2 are symmetric, then so is $\sigma_1 * \sigma_2$. Furthermore, if σ_1 and σ_2 are convex, then so is $\sigma_1 * \sigma_2$. The truth of the last statement is implied by the treatment of a problem on rearrangements (cf. Hardy-Littlewood-Pólya [7], pp. 273-274); a direct proof may be found in § 4, where the theorem is extended to the multi-dimensional case (Theorem XIII). The truth of the statement is almost trivial in view of the statistical interpretation of the convolution process, mentioned above (cf. § 3). Examples show that if $\sigma_1 * \sigma_2$ and σ_2 are convex, then σ_1 need not be convex; cf., e. g., the convex distribution function occurring in Theorem V.

The Fourier transform of a distribution function $\sigma(x)$ is defined as the Stieltjes integral

$$(1) \qquad L(t) = L(t;\sigma) = \int_{-\infty}^{+\infty} e^{itx} d\sigma(x), \text{ where } -\infty < t < +\infty.$$

It is easy to see that the function $L(t;\sigma)$ is uniformly continuous in the

infinite interval $-\infty < t < +\infty$; it may be nowhere absolutely continuous, since the Weierstrass example

$$(1-a)\sum_{k=0}^{\infty} a^k \exp(ib^k t)$$

is, for suitable values of a and b, an $L(t;\sigma)$ which is nowhere differentiable. There belongs to every $L(t;\sigma)$ but one $\sigma(x)$, since

$$\sigma(x) = \sigma(0) + (2\pi i)^{-1}\int_{-\infty}^{+\infty} t^{-1}(1-e^{-itx})L(t;\sigma)\,dt, \text{ where } \int_{-\infty}^{+\infty} = \lim_{T=+\infty}\int_{-T}^{T}.$$

This is Lévy's inversion formula (cf., e. g., Wiener [21], Theorem 36 or Haviland [10]). It implies that if the integral of $|L(t;\sigma)|$ over $-\infty < t < +\infty$ is finite, then $\sigma(x)$ has for $-\infty < x < +\infty$ a uniformly continuous and bounded derivative which may be obtained by formal differentiation,

$$(2) \qquad \sigma'(x) = (2\pi)^{-1}\int_{-\infty}^{+\infty} e^{-ixt}L(t;\sigma)\,dt.$$

If there exist two positive constants, a and λ, such that

$$(3) \qquad L(t;\sigma) = O(\exp\{-a\,|\,t\,|^{\lambda}\}) \text{ as } t\to\pm\infty,$$

then all derivatives of $\sigma(x)$ exist for every x and may be obtained by successive formal differentiation of (2). If in particular $\lambda = 1$, then there exists in a strip $-\alpha < y < \alpha$ of the z-plane, where $z = x + iy$, a regular and bounded function which becomes the distribution function $\sigma(x)$ if $y = 0$, the least upper bound of the admissible values of the width 2α being not less than $2a$; cf. Wintner [25]. Hence if $\lambda > 1$, then one may choose 2α arbitrarily large, and so σ is an entire function which is bounded in every strip parallel to the real axis. It may be mentioned that σ cannot be a rational entire function, since $\sigma(-\infty) = 0$ and $\sigma(+\infty) = 1$.

It is obvious from (1) that $L(-t;\sigma) = L(t;\bar{\sigma})$, where $\bar{\sigma}(x) = 1 - \sigma(-x)$. Furthermore, $L(t;\sigma)$ and $L(-t;\sigma)$ are conjugated complex values. Since $\sigma(x)$ is monotone non-decreasing and has the total variation 1, it is also clear from (1) that $L(0;\sigma) = 1$ and that $|\,L(t;\sigma)\,| \leqq 1$ for every t. Suppose that σ is such that $|\,L(t;\sigma)\,| = 1$ holds not only for $t = 0$ but for at least one $t = t_0 \neq 0$ also. Then $L(0;\sigma) - e^{i\vartheta}L(t_0;\sigma) = 0$ for some real $\vartheta = \vartheta(t_0)$; hence, on taking the real part,

$$\int_{-\infty}^{+\infty} f(x)\,d\sigma(x) = 0, \text{ where } f(x) = 1 - \cos(\vartheta + t_0 x).$$

Consequently, since $f(x) \geqq 0$ is continuous, $f(x) = 0$ at every point x of the spectrum of σ. Now every root of the equation $f(x) = 0$ satisfies the congruence $t_0 x \equiv -\vartheta \pmod{2\pi}$, where $t_0 \neq 0$ and $\vartheta = \vartheta(t_0)$. Hence in order that $|L(t;\sigma)| = 1$ holds for at least one $t \neq 0$, it is necessary that the spectrum of σ be a sequence contained in an arithmetical progression. This condition (cf. Cannon and Wintner [2]) is sufficient also, since if it is satisfied, $L(t;\sigma)$ is a periodic function and must, therefore, attain the value attained at $t = 0$ not only at $t = 0$. If $|L(t;\sigma)| = 1$ for every t, i.e., if t_0 is arbitrary, then the spectrum must be contained in an arithmetical progression of arbitrary difference, hence it must consist of a single point $x = b$. The distribution function is then $\chi(x - b)$ which has Fourier transform e^{itb}. In particular, $L(t;\sigma) \equiv 1$ belongs to $\sigma = \chi$.

The Fourier transform of the distribution function $\sigma(ax)$ is $L(t/a;\sigma)$. If $\sigma(x) = \sigma(ax)$, where $a > 0$, then $a = 1$ unless $\sigma(x) = \chi(x)$. This necessary condition is sufficient also, i.e., the distribution functions which are similar to $\chi(x)$ are identical with $\chi(x)$.

For a given distribution function σ, put $[\sigma] = \limsup |L(t;\sigma)|$, where $t \to +\infty$ or $t \to -\infty$. It is clear from $|L(t;\sigma)| \leqq 1$ that $0 \leqq [\sigma] \leqq 1$. If σ is absolutely continuous, then $[\sigma] = 0$ in view of the Riemann-Lebesgue lemma. If $[\sigma] = 0$, then σ need not be absolutely continuous. If $[\sigma] = 0$, then $\sigma(x)$ is everywhere continuous. For if σ has at least one discontinuity point, then $L(t;\sigma)$ has a component which is almost periodic in the sense of Bohr, is not identically zero and is not destroyed by the complementary component of $L(t;\sigma)$; in this connection cf. Haviland and Wintner [11]. If σ is purely discontinuous, then $L(t;\sigma)$ is almost periodic in the sense of Bohr, and so $[\sigma] = 1$ in view of $L(0;\sigma) = 1$. It would be interesting to know whether $[\sigma] = 1$ may or may not hold if σ is not purely discontinuous. It may be mentioned that there exists for $\epsilon > 0$ a distribution function σ which is nowhere discontinuous but such that $[\sigma] > 1 - \epsilon$. First, there exists for every positive $a < \frac{1}{2}$ a distribution function σ which is continuous but not absolutely continuous and is defined by the condition

$$L(t;\sigma) = \prod_{k=1}^{\infty} \cos(a^k t);$$

cf. Jessen and Wintner [12], § 6 and Kershner and Wintner [13]. Now choose $a = m^{-1}$, where $m > 2$ is a fixed integer, and put $t = 2\pi m^j$, where $j = \pm 1, \pm 2, \cdots$. Then

$$L(2\pi m^j;\sigma) = \prod_{k=1}^{\infty} \cos(m^{-k} 2\pi m^j) = \prod_{k=1}^{\infty} \cos(2\pi/m^k)$$

for every j. Hence, on letting $|j| \to \infty$ and keeping $a = m^{-1}$ fixed,

$$[\sigma] \geqq \prod_{k=1}^{\infty} |\cos(2\pi/m^k)|,$$

where σ depends on m. The last product clearly tends to 1 as $m \to \infty$, showing that $[\sigma] > 1 - \epsilon$ for a suitable continuous σ.

One of the reasons for the importance of the Fourier transformation in the theory of distribution functions is the fact that $L(t; \sigma_1 * \sigma_2)$ is more easily obtained than $\sigma_1 * \sigma_2$. In fact,

$$L(t; \sigma_1 * \sigma_2) = L(t; \sigma_1) L(t; \sigma_2); \tag{4}$$

cf. Haviland [9], [10].

A sequence $\{\sigma_n\}$ of distribution functions is convergent if and only if the sequence $\{L(t; \sigma_n)\}$ of the Fourier transforms is uniformly convergent in every fixed finite interval $|t| <$ const., and $\lim L(t; \sigma_n)$ is then $L(t; \lim \sigma_n)$. This is Lévy's "continuity theorem" for the Fourier transforms of distribution functions. The theorem implies an existence statement, namely the assertion that the distribution function $\lim \sigma_n$ exists. A simple proof has recently been given by the present author; cf. Haviland [10], where further references are given.

The infinite convolution $\sigma_1 * \sigma_2 * \cdots$ is said to converge to the distribution function σ if $\sigma_1 * \cdots * \sigma_n \to \sigma$ as $n \to \infty$. Since

$$L(t; \sigma_1 * \cdots * \sigma_n) = L(t; \sigma_1) \cdots L(t; \sigma_n)$$

in view of (4), the continuity theorem implies that $\sigma_1 * \sigma_2 * \cdots$ is a convergent infinite convolution if and only if the infinite product

$$\prod_{k=1}^{\infty} L(t; \sigma_k) \tag{5}$$

is uniformly convergent in every fixed finite interval $|t| <$ const., and that this product is then $L(t; \sigma_1 * \sigma_2 * \cdots)$. For a detailed theory of infinite convolutions cf. Jessen and Wintner [12]. The references given there must be completed by one to the paper [8] of Hausdorff who was apparently the first to consider an infinite convolution in a particular case (cf. § 2 below).

Since two similar distribution functions differ from each other but in the choice of the unit of x, it is clear from the statistical meaning of the convolution process that if σ is a possible universal law of random distribution, then the convolution of any pair of distribution functions which are similar

to σ must be again similar to σ; cf. Bessel [1], Cauchy [4], Pólya [17]. If this condition is satisfied, σ is termed a stable distribution function. Thus the condition of stability is that there exists for every $a > 0$ and every $b > 0$ a $c = c(a, b) > 0$ such that

$$\sigma(x/a) * \sigma(x/b) = \sigma(x/c),$$

i. e.,

$$L(at; \sigma) L(bt; \sigma) = L(ct; \sigma). \tag{6}$$

This does not imply that there exists for every $a > 0$ and every $c > 0$ a $b = b(a, c) > 0$. It is clear from a remark made above with regard to similar distribution functions that if σ is stable, c is uniquely determined by a and b unless $\sigma = \chi$.

2. *Fourier transforms of convex analytic distributions.* The kernel of the representation of the function

$$1/\operatorname{Cosh} t \tag{7}$$

as a Fourier cosine transform is (i) positive and (ii) monotone decreasing. This is implied by the fact that (7) is, up to constant factors, self-reciprocal with respect to the Fourier cosine transform. Theorem I will deal with a large class of Fourier transforms satisfying conditions (i) and (ii). The class in question contains each of the functions

$$(it)^\nu / J_\nu(it), \text{ where } \nu > -1. \tag{8}$$

The example (8) implies (7), since $\nu = -\frac{1}{2}$ is not excluded; for $\nu = \frac{1}{2}$ one obtains $t/\operatorname{Sinh} t$. The treatment of the general function class in question will be based on an infinite convolution, considered by Hausdorff [8] in the explicitly available case (7). For the treatment of the general case, which is not self-reciprocal, an existence theorem is needed. Although a theorem to this effect might easily be proven directly, this will not be done, since the existence statement in question is implied by the continuity theorem mentioned in § 1.

Let $f(z)$ be an even transcendental entire function which has but real zeros and is positive at $z = 0$. Suppose further that the entire function $f(\sqrt{z})$ is of genus 0. Thus

$$A = \sum_{k=1}^{\infty} a_k^2 < +\infty, \qquad (a_1 \geqq a_2 \geqq \cdots; a_k > 0), \tag{9}$$

where

$$\pm 1/a_1, \pm 1/a_2, \cdots \tag{10}$$

denote the zeros of $f(z) = f(-z)$, and

$$f(z) = f(0) \prod_{k=1}^{\infty} (1 - a_k^2 z^2), \tag{11}$$

where $f(0) > 0$. In particular,

$$\max_{|z| \leqq |t|} | f(z) | = f(it), \tag{12}$$

where $-\infty < t < +\infty$. Examples of functions $f(z)$ which satisfy the conditions just mentioned are infinite products like

$$f(z) = \prod_{n=1}^{\infty} \cos (b_n z) \text{ and } f(z) = \prod_{n=1}^{\infty} J_0(b_n z); \ (\sum_{n=1}^{\infty} b_n^2 < +\infty; \ b_n > 0),$$

(occurring in connection with the simplest types of infinite convolutions), the trigonometrical integrals

$$f(z) = \int_0^1 \psi(x) \cos zx dx$$

investigated by Pólya [16] and, in particular, the functions

$$f(z) = z^{-\nu} J_\nu(z), \text{ where } \nu \geqq -\tfrac{1}{2}.$$

All these examples $f(z)$ admit of a representation of the form

$$f(z) = \int_0^{+\infty} \cos zx d\phi(x), \tag{13}$$

where $\phi(x)$ is monotone, bounded and not everywhere constant. On the other hand, not every $f(z)$ under consideration is representable in the form (13). In fact, (13) implies (12), hence it also implies

$$\max_{|z| \leqq |t|} | f(z) | > C e^{c|t|}, \tag{14}$$

where C and c are positive constants. Now suppose that the zeros (10) of $f(z)$ are so scarce as to make the series $a_1 + a_2 + \cdots$ convergent. Then

$$h(z) = \prod_{k=1}^{\infty} (1 - a_k z)$$

is a canonical product of order 0. Since $f(z) = f(0) h(z) h(-z)$ in view of (11), the function $f(z)$ also is of order 0, hence cannot satisfy the condition (14), which is a necessary condition for (13). In the following existence theorem it is not assumed that $f(z)$ is representable in the form (13).

THEOREM I. *Let $f(z)$, where $f(0) > 0$, be an even transcendental entire*

function which is of genus 0 in z^2 and has but real zeros. Then there exists a distribution function $\sigma(x)$ which is convex in the sense defined in § 1, has for $-\infty < x < +\infty$ bounded derivatives of arbitrarily high order and is connected with f by the formula

$$f(0)/f(it) = L(t;\sigma); \text{ where } -\infty < t < +\infty, \tag{15}$$

L being the Fourier transform (1); *and so* $L(t;\sigma) = 2\int_0^{+\infty} \cos tx d\sigma(x)$ *in view of the symmetry of* σ.

If $f(z)$ *satisfies the additional condition* (14), *then the distribution function* $\sigma(x)$ *implicitly defined by* (15) *is regular and bounded in a strip* $-\alpha < y < \alpha$ *of the z-plane, where* $z = x + iy$.

Remark. The example $f(z) = \cos z$ shows that σ need not be an entire function. In fact, (7) is, up to constant factors, self-reciprocal and has therefore poles on the boundary of a strip of finite width. Also, the function $f(0)/f(iz)$ represented for real z by (15) always has on the imaginary axis a pole in a finite distance from $z = 0$.

Proof. Let $\gamma = \gamma(x)$ denote distribution function

$$\gamma(x) = \tfrac{1}{2}\int_{-\infty}^{x} e^{-|y|}dy.$$

Since the derivative $\gamma'(x) = \gamma'(-x)$ is a decreasing function of $|x|$, the distribution function γ is convex in the sense of § 1. Furthermore,

$$L(t;\gamma) = \int_{-\infty}^{+\infty} e^{itx}\,\tfrac{1}{2}e^{-|x|}\,dx = \int_0^{+\infty} \cos tx\, e^{-x}\,dx,$$

i. e., $L(t;\gamma) = (1+t^2)^{-1}$. Hence on denoting by $\sigma_k(x)$ the distribution function $\gamma(x/a_k)$,

$$L(t;\sigma_k) = L(a_k t;\gamma) = (1 + a_k{}^2t^2)^{-1},$$

and so the product (5) is, in view of (9), uniformly convergent in every fixed finite interval $|t| < \text{const.}$ Consequently, there exists a distribution function σ represented by the infinite convolution $\sigma_1 * \sigma_2 * \cdots$, and

$$L(t;\sigma) = \prod_{k=1}^{\infty} L(t;\sigma_k) = \prod_{k=1}^{\infty}(1 + a_k{}^2t^2)^{-1},$$

which proves (15) in view of (11). Since $a_k > 0$,

$$L(t;\sigma) = [\prod_{k=1}^{\infty}(1 + a_k{}^2t^2)]^{-1} = O(|t|^{-N}),\ t \to \pm\infty,$$

holds for every fixed N. Hence $\sigma(x)$ has for $-\infty < x < +\infty$ bounded derivatives of arbitrarily high order. This follows from the last estimate, where N may be arbitrarily large, by successive differentiation of (2). Since $\gamma(x)$ is convex, so is $\gamma(x/a_k) = \sigma_k(x)$ for every k; hence $\sigma_1 * \cdots * \sigma_m$ and therefore

$$\lim_{m=\infty} \sigma_1 * \cdots * \sigma_m = \sigma_1 * \sigma_2 * \cdots = \sigma$$

also is convex. Finally, if $f(z)$ satisfies the additional condition (14), then it is clear from (12) and (15) that (3) is satisfied for $\lambda = 1$ and $a = c > 0$. This completes the proof of Theorem I.

Since σ is convex, hence symmetric,

$$\int_{-\infty}^{+\infty} f(x)\, d\sigma(x) = 2\int_{0}^{+\infty} f(x)\, d\sigma(x) \quad \text{or} \quad = 0$$

according as $f(x)$ is even or odd, provided that the latter integral is convergent. Information about the behavior of σ at $x = +\infty$, hence at $x = -\infty$, is given by

THEOREM II. *Let $f(z)$ satisfy the general requirements of Theorem* I *but not necessarily the additional condition* (14), *and let σ be the distribution function defined by* (15). *Then all integrals*

$$\int_{0}^{+\infty} x^n d\sigma(x) \qquad (n = 0, 1, 2, \cdots) \tag{16}$$

are convergent and belong to a determined Stieltjes moment problem. Furthermore,

$$\int_{x}^{+\infty} d\sigma(y) = O\,(\exp\{-cx^{\frac{4}{3}}\}) \text{ as } x \to +\infty, \tag{17}$$

where $c > 0$.

Remark. The estimate leading to (17) is so crude that it does not attach any particular significance to the numerical value $\frac{4}{3}$ of the exponent of x. On the other hand, (17) describes the true situation up to the numerical value of this exponent. For if $f(z) = \cos z$, it is clear from the example (7) that (17) becomes false if one replaces $\frac{4}{3}$ by $1 + \epsilon$. Thus the best value of the exponent is somewhere between $\frac{4}{3}$ and 1, and it is not proven that it is less than 1. That it cannot be greater than 1, agrees with the fact that $1/f(z)$ is not an entire function.

Proof. Put

$$\mu_r(\psi) = \int_{-\infty}^{+\infty} x^r d\psi(x), \qquad (r = 1, 2, \cdots),$$

and let γ and σ_k denote the distribution functions defined in the Proof of Theorem I. Thus $\mu_{2m+1}(\gamma) = 0$, while

$$\mu_{2m}(\gamma) = \int_{-\infty}^{+\infty} x^{2m} \tfrac{1}{2} e^{-|x|} dx = (2m)\,!;$$

hence, since $\sigma_k(x) = \gamma(x/a_k)$,

$$\mu_{2m}(\sigma_k) = a_k{}^{2m} \mu_{2m}(\gamma) = (2m)\,!\,a_k{}^{2m}.$$

Since, by Stirling's formula, $N\,! < (\alpha N)^N$ for every positive integer N and for some constant $\alpha > 0$, it follows that

$$\mu_{2m}(\sigma_k) < (2\alpha m\, a_k)^{2m}.$$

Now

$$\mu_{2m}(\sigma_1 * \cdots * \sigma_n) \leqq m^m \Sigma\, C_{hj\ldots}\, \mu_{2h}(\sigma_1)\mu_{2j}(\sigma_2)\cdots,$$

where $C_{hj\ldots}$ denotes the multinomial coefficient

$$C_{hj\ldots} = (h + j + \cdots)\,!/(h\,!\,j\,!\cdots)$$

and the summation Σ runs through all collections of n non-negative integers $h, j, \cdots$ for which $h + j + \cdots = m$; for proof cf. Wintner [28], where reference is given to a similar inequality of Paley and Zygmund. Thus

$$\mu_{2m}(\sigma_1 * \cdots * \sigma_n) < m^m \Sigma\, C_{hj\ldots}\, (2\alpha\, h\, a_1)^{2h}(2\alpha\, j\, a_2)^{2j}\cdots,$$

or, since $h + j + \cdots = m$,

$$\mu_{2m}(\sigma_1 * \cdots * \sigma_n) < m^m (4\alpha^2)^m \Sigma\, C_{hj\ldots}\, a_1{}^{2h} a_2{}^{2j}\cdots (h^h j^j \cdots)^2.$$

On combining this inequality with the crude estimate

$$h^h j^j \cdots \leqq m^h m^j \cdots = m^m$$

and with the multinomial theorem, i. e., with the identity

$$\Sigma\, C_{hj\ldots}\, a_1{}^{2h} a_2{}^{2j} \cdots = (a_1{}^2 + \cdots + a_n{}^2)^m,$$

it is seen that

$$\mu_{2m}(\sigma_1 * \cdots * \sigma_n) < m^m (4\alpha^2)^m (a_1{}^2 + \cdots + a_n{}^2)^m (m^m)^2.$$

Consequently, from (9),

$$\mu_{2m}(\sigma_1 * \cdots * \sigma_n) < C^m m^{3m},$$

where $C = 4\alpha^2 A$ is independent both of n and m. On placing $\rho_n = \sigma_1 * \cdots * \sigma_n$ and letting $n \to \infty$ for a fixed m, Helly's theorem on term-by-term integration shows that

$$\int_{-R}^{R} x^{2m} d\rho_n(x) \to \int_{-R}^{R} x^{2m} d\sigma(x)$$

for every fixed $R > 0$; for $\rho_n = \sigma_1 * \cdots * \sigma_n$ tends, as $n \to \infty$, to the infinite convolution $\sigma_1 * \sigma_2 * \cdots$ which defined σ in the Proof of Theorem I. On letting $R \to +\infty$ for a fixed m, it follows that

$$\mu_{2m}(\sigma) \leqq C^m m^{3m} \tag{18}$$

for every m. Now if M_n denotes the integral (16) and if every M_{2m} is convergent, then so is every M_{2m+1} in view of the Schwarz inequality; and every M_{2m} is convergent in virtue of (18), since $\mu_{2m} = 2M_{2m}$. It also follows from (18) that $(M_{2m})^{-1/(4m)} \geqq \text{const.}/m^{3/4}$, which implies the divergence of the series

$$\sum_{m=1}^{\infty} (M_m)^{-1/(2m)}.$$

Hence the criterion of Carleman [3], p. 81, for the Stieltjes case shows that (16) belongs to a determined Stieltjes moment problem.

The transition from (18) to (17) requires but a standard argument (cf. Zygmund [30], p. 124). In fact, since

$$\int_{0}^{+\infty} x^{2m} d\sigma(x) = \tfrac{1}{2}\mu_{2m}(\sigma) \text{ and } \int_{0}^{+\infty} d\sigma(x) = \tfrac{1}{2},$$

Hölder's inequality shows that

$$\int_{0}^{+\infty} x^{2m/3} d\sigma(x) \leqq [\int_{0}^{+\infty} (x^{2m/3})^3 d\sigma(x)]^{1/3} [\int_{0}^{+\infty} 1^{3/2} d\sigma(x)]^{2/3} < [\mu_{2m}(\sigma)]^{1/3}.$$

Thus it is seen from (18) that

$$\int_{0}^{+\infty} x^{2m/3} d\sigma(x) < (Bm)^m,$$

where $B = C^{1/3}$ and $m = 1, 2, \cdots$. Hence it is clear from Stirling's formula,

$$m!^{-1}(mB)^m \sim (Be)^m/(2\pi m)^{1/2}, \quad m \to \infty,$$

that the power series

$$p(z) = \sum_{m=0}^{\infty} m!^{-1} \int_{0}^{+\infty} x^{2m/3} d\sigma(x) \cdot z^m$$

has a non-vanishing radius of convergence. Since the coefficients of this power series are positive, it follows by writing y instead of x that for every sufficiently small $z = c > 0$

$$\int_0^{+\infty} \exp(cy^{2/3})\,d\sigma(y) = p(c) < +\infty.$$

Consequently, if $x > 0$,

$$\exp(cx^{2/3}) \int_x^{+\infty} d\sigma(y) \leqq \int_x^{+\infty} \exp(cy^{2/3})\,d\sigma(y) \leqq p(c) = \text{const.}$$

This completes the proof of Theorem II.

Let Ξ be defined by $\Xi(z) = \xi(\tfrac{1}{2} + iz)$, where $\xi(z)$ denotes the Riemann ξ-function; cf. Pólya [19], Titchmarsh [20], p. 43.

THEOREM III. *On Riemann's hypothesis,*

$$\Xi(0)/\Xi(it) = L(t;\sigma), \quad -\infty < t < +\infty,$$

where $\sigma(x)$ is a convex distribution function which is regular and bounded in a strip $-\alpha < y < \alpha$.

Proof. According to the Hadamard theory, the entire function $\Xi(\sqrt{z})$ is of genus 0 in z. Furthermore, the Riemann integral representation of $\Xi(z)$ is of the form (13), where $\phi'(x) > 0$. Finally, Riemann's hypothesis is that all zeros of $\Xi(z)$ are real. Hence Theorem III is implied by Theorem I.

THEOREM IV. *Independently of the Riemann hypothesis,*

$$\Xi(t)/\Xi(0) = L(t;\sigma), \quad -\infty < t < +\infty,$$

where $\sigma(x)$ is a convex distribution function which is regular and bounded in a strip $-\alpha < y < \alpha$.

For the proof of the convexity of σ cf. Wintner [27]. The Riemann integral representation of Ξ shows that the least upper bound of α is $\pi/8$ and that the lines $y = \pm\pi/8$ form a natural boundary of σ. The point in Theorem IV is that σ is convex. This implies the following result which is less deep and is of an older date, since it has been pointed out in 1916 in an equivalent form by Wilton [22].

COROLLARY. *On denoting by $Z(s)$ the meromorphic function $\zeta(s)\Gamma(\tfrac{1}{2}s)\pi^{-s/2}$,*

$$Z(\tfrac{1}{2} + it)/Z(\tfrac{1}{2}) = L(t;\rho), \quad -\infty < t < +\infty,$$

where ρ is a convex distribution function.

This is a consequence of Theorem IV, but not conversely. First,

$$\Xi(t) = -\tfrac{1}{2}(\tfrac{1}{4} + t^2)Z(\tfrac{1}{2} + it)$$

by the Riemann definition of $\Xi(t) = \xi(\frac{1}{2} + it)$; cf. Titchmarsh [20], p. 43. Put $\delta(x) = \gamma(\frac{1}{2}x)$, where $\gamma(x)$ is the convex distribution function occurring in the proof of Theorem I, so that $L(t;\gamma) = (1 + t^2)^{-1}$. Thus δ is a convex distribution function. Hence, if σ denotes the distribution function occurring in Theorem IV, then $\rho = \delta * \sigma$ also is convex. Now, on using (4),

$$L(t;\rho) = L(t;\delta)L(t;\sigma) = L(2t;\gamma)\Xi(t)/\Xi(0),$$

where $L(2t;\gamma) = (1 + (2t)^2)^{-1}$; hence

$$L(t;\rho) = (1 + 4t^2)^{-1}[-\tfrac{1}{8}(1 + 4t^2)Z(\tfrac{1}{2} + it)]/\Xi(0) = -\tfrac{1}{8}Z(\tfrac{1}{2} + it)/\Xi(0).$$

Consequently,

$$Z(\tfrac{1}{2} + it) = Z(\tfrac{1}{2})L(t;\rho),$$

where ρ is convex; q. e. d. It may be mentioned that the Fourier transform $L(t;\rho)$ is precisely the one which occurs in Hardy's proof for the existence of infinitely many real zeros of $\zeta(\frac{1}{2} + it)$.

The absolutely convergent product (19) considered in the next Theorem is often mentioned in the elementary theory of canonical products; cf., e. g., Francis and Littlewood [5], p. 5. Although the zeros of the entire function (19), which is of genus 0 in t^2, are all real and equidistant, the rapid increase of the multiplicity of these zeros clearly puts the function beyond Pringsheim's "normal type of order 1." Apparently it is not known whether or not the function is related to solutions of standard linear differential equations of the second order.

THEOREM V. *There exists a convex distribution function τ such that*

$$\prod_{k=1}^{\infty} \cos(t/k) = L(t;\tau), \tag{19}$$

and $\tau(x)$ is regular and bounded in a strip $-\alpha < y < \alpha$. Furthermore, there exists a $c > 0$ such that

$$\tau(x) = 1 - \tau(-x) = O(\exp\{-cx^2\}) \quad \text{as } x \to -\infty$$

and Carleman's series

$$\sum_{m=1}^{\infty} (\mu_{2m})^{-1/(2m)}, \quad \text{where} \quad \mu_{2m} = \int_{-\infty}^{+\infty} x^{2m} d\tau(x) < +\infty,$$

is divergent.

For proof cf. Wintner [28], and [29], p. 838. The least upper bound of the admissible values of the width 2α is not less than 1; cf. Jessen and Wintner [12], p. 62. It is not known if the boundary of the widest strip is a natural boundary of τ. Nor is it known that the widest strip has a boundary; it is not even proven that τ is not an entire function. A simple proof for the existence of an $\alpha > 0$ proceeds as follows. Let

$$a_1 > 0, a_2 > 0, \cdots \text{ and } a_1^2 + a_2^2 + \cdots < +\infty$$

and let k be so large that $a_k \mid t \mid < 1$, where $\mid t \mid \neq 0$ is fixed. Since there exist a positive constant $C < 1$ and a positive constant B such that

$$0 < \cos\vartheta < 1 - C\vartheta^2 \text{ and } \log(1-\vartheta) < -B\vartheta, \text{ where } 0 < \vartheta < 1,$$

it is clear that

$$0 < \prod_{a_k|t|<1} \cos(a_k t) < \prod_{a_k|t|<1} (1 - Ca_k^2 t^2) < \exp \sum_{a_k|t|<1} - BCa_k^2 t^2;$$

hence

$$(20) \qquad |\prod_{k=1}^{\infty} \cos(a_k t)| \leqq \prod_{a_k|t|<1} \cos(a_k t) = \exp O(-At^2 \sum_{a_k|t|<1} a_k^2),$$

where $A = BC > 0$ and $t \to \pm\infty$. On choosing $a_k = k^{-1}$, it follows from (20) that (19) satisfies (3) with $\lambda = 1$. As far as the convexity of τ is concerned, cf. (35) below.

Remark. It is instructive to contrast the distribution functions σ and τ occurring in Theorems IV and V. The relation (19) means that $\tau(x)$ is the convolution of the infinite sequence $\beta(x), \beta(2x), \beta(3x), \cdots$ of Bernoulli distribution functions, where $\beta(x) = 0, \frac{1}{2}$ or 1 according as x lies on the left, in the interior or on the right of the interval $-1 < x < 1$. This *statistical factorization* of τ yields for the entire function $L(t;\tau)$ the factorization (19) which, being *not* the canonical factorization of Weierstrass-Hadamard, puts the reality of all zeros of $L(t;\tau)$ into evidence. In Theorem IV, the reality of all zeros of $L(t;\sigma)$ is Riemann's hypothesis. Thus one should like to obtain instead of the Weierstrass-Hadamard factorization of $L(t;\sigma)$ a factorization (5) of $L(t;\sigma)$ into a product of Fourier transforms $L(t;\sigma_k)$ or, what is the same thing, a statistical decomposition of σ into an infinite convolution $\sigma_1 * \sigma_2 * \cdots$, based on particular statistical and not on general function-theoretical properties. A quantitative illustration of the possible analogy between $L(t;\sigma)$ and $L(t;\tau)$ may be obtained as follows. Let $N_0(T)$ be the number of zeros of $L(t;\sigma) = \Xi(t)/\Xi(0)$ in the interval $0 < t < T$ and let $N'(t)$ denote the corresponding function belonging to $L(t;\tau)$, each

zero being counted according to its multiplicity. The Riemann-Mangoldt asymptotic formula implies that $N_0(T) \sim (2\pi)^{-1}T \log T$ on Riemann's hypothesis. This relation, although so far unproven, means very much less than Riemann's hypothesis and is possibly the estimate which Riemann had in mind in his famous statement: "Man findet . . . *etwa* . . . so viel reelle Wurzeln . . ." (the italics are mine). On the other hand, $N'(T)$ is the number of zeros of

$$\prod_{k=1}^{[2T/\pi]} \cos(t/k)$$

between $t = 0$ and $t = T$; for if $k > 2T/\pi$, then $T/k < \frac{1}{2}\pi$; and so the factor $\cos(t/k)$ in (19) does not contribute to $N'(T)$. Since $\cos t$ has between $t = 0$ and $t = T$ about $\pi^{-1}T$ zeros, it follows that

$$N'(T) \sim \sum_{k=1}^{[2T/\pi]} \pi^{-1}T/k \sim \pi^{-1}T \log T.$$

Hence $N_0(T) \sim \frac{1}{2}N'(T)$ on Riemann's hypothesis. The distance between two subsequent discontinuity points of $N_0(T)$ tends, according to Littlewood, on Riemann's hypothesis to zero (cf. Titchmarsh [20], p. 60), while the distance between two subsequent discontinuity points of $N'(T)$ has the constant value π. Thus $N_0(T) \sim \frac{1}{2}N'(T)$ agrees with the rapid increase of the multiplicities of the zeros of $L(t;\tau)$.

If p_k denotes the k-th prime number,

$$\prod_{k=1}^{\infty} \cos(t/p_k) \tag{21}$$

also is the Fourier transform L of a distribution function which has derivatives of arbitrarily high order along the real axis; cf. Wintner [26]. The latter statement is clear from (20) also. In fact, $a_k = p_k^{-1}$, so that (20) implies that (3) is satisfied for every $\lambda < 1$. It is not known whether or not the distribution function belonging to (21) is analytic,[1] and the question of convexity also is unanswered.

[1] On the other hand, it is easy to prove that the infinite convolution in question is *a distribution function $\sigma(x)$ which is, for $-\infty < x < +\infty$, quasi-analytic in the sense of Denjoy and Carleman.* First,

$$L(t;\sigma) = O(\exp - At^2 \sum_{p_n > |t|} p_n^{-2}), \qquad t \to \pm\infty, \quad A > 0,$$

in view of (20). Hence it is clear from the prime-number theorem ($p_n \sim n \log n$) or even from the elementary inequalities of Tchebycheff that

$$L(t;\sigma) = O(\exp - C\,|t|/\log|t|), \qquad t \to \pm\infty, \tag{I}$$

The simplest example of Theorem I was the standard formula expressing the fact that (7) is, up to constant factors, a self-reciprocal function of the Fourier cosine transform. Correspondingly, the formula

$$(22)\qquad |\Gamma(b+\tfrac{1}{2}it)|^2 = C_b \int_{-\infty}^{+\infty} (\operatorname{Cosh} x)^{-2b} e^{itx} dx;\ b>0,\ C_b = 2^{1-2b}\Gamma(2b),$$

which generalizes the previous one from $b = \frac{1}{2}$ to an arbitrary $b > 0$ and also is standard (cf., e. g., Mathias [15], p. 113), is but the simplest illustration of a general theorem:

THEOREM VI. *If the entire function $g(z)$ is real along the real axis, has infinitely many zeros which are all zero or negative and is of genus 0 or 1, then there exists for every $b > 0$ a convex distribution function $\sigma = \sigma_b = \sigma_b(x)$ such that*

$$(23)\qquad |g(b+it)|^{-2} = |g(b)|^{-2} L(t; \sigma_b);\ -\infty < t < +\infty.$$

Furthermore, this distribution function has derivatives of arbitrarily high order for $-\infty < x < +\infty$ and its behavior at $x = \infty$ is the same as in Theorem II. *If in addition the zeros of $g(z)$ are so abundant and so regularly situated as to imply a certain inequality analogous to the additional condition (14) of Theorem* I, *then σ_b is regular and bounded in a strip $-\alpha < y < \alpha$.*

Remark. If

$$g(z) = 1/\Gamma(z) = ze^{Cz} \prod_{m=1}^{\infty} (1 + z/m) e^{-z/m};$$

the genus of $g(z)$ is 1 and the zeros $z = -m < 0$ are sufficiently abundant and "regularly distributed." Thus (22), where $g(z) = 1/\Gamma(\frac{1}{2}z)$, is an explicit example of Theorem VI. In fact, $(\operatorname{Cosh} x)^{-2b}$ is for every $b > 0$ a positive decreasing function of $|x|$. Since $(\operatorname{Cosh} z)^{-2b}$ has on the boundary of a strip $-\alpha < y < \alpha$ of width $2\alpha = \pi$ poles $(2b = 1, 2, \cdots)$ or logarithmical singularities $(2b \neq 1, 2, \cdots)$, the example (22) also shows that the last statement

---

where C is a positive constant. Now it is easy to see that

$$(\text{II})\qquad \int_{e}^{+\infty} t^{n-1} \exp(-Ct/\log t)\, dt < (\text{const.}\ n \log n)^n.$$

This follows from Stirling's formula. According to a remark of Mr. R. B. Kershner, one obtains a short proof of (II) by introducing instead of t the integration variable $t/\log t$. Now it is obvious from (I) and (II) by successive differentiation of (2) that

$$|\sigma^{(n)}(x)| < (Mn \log n)^n,$$

where M is independent both of n and x. Since $\Sigma(n \log n)^{-1}$ is a divergent series, the quasi-analyticity of $\sigma(x)$ follows by the Denjoy-Carleman criterion; cf. Carleman [3].

of Theorem VI cannot be improved to the statement that σ_b is an entire function.

Proof. If Theorem VI is true on the assumption $g(0) \neq 0$, it is true also when $g(0) = 0$. For if $g(z)$ vanishes at $z = 0$ in the order $d \geqq 1$, there exists, by assumption, a convex distribution function σ_b such that

$$| z^{-d}g(z) | = \text{const. } L(t;\sigma_b), \text{ where } z = b + it.$$

Thus, since $L(0;\sigma_b) = 1$,

$$| g(b+it) |^{-2} = | g(b) |^{-2} L(t;\sigma_b)(1 + t^2/b^2)^{-d},$$

where $g(b) \neq 0$, since $b > 0$. Now if $\gamma(x)$ denotes the convex distribution function occurring in the Proof of Theorem I and if $\gamma_b(x)$ is the similar distribution function $\gamma(bx)$, then $L(t;\gamma) = (1 + t^2)^{-1}$, hence

$$(1 + t^2/b^2)^{-d} = (L(t/b;\gamma))^d = (L(t;\gamma_b))^d = L(t;\gamma_{bd}),$$

where γ_{bd} denotes the distribution function $\gamma_b * \cdots * \gamma_b$ (d times). Thus γ_{bd} is a convex distribution function and

$$| g(b+it) |^{-2} = | g(b) |^{-2} L(t;\sigma_b) L(t;\gamma_{bd}).$$

Now $L(t;\sigma_b)L(t;\gamma_{bd}) = L(t;\sigma_b * \gamma_{bd})$, and $\sigma_b * \gamma_{bd}$ also is a convex distribution function. Consequently, the case $g(0) = 0$ is reducible to the case $g(0) \neq 0$.

Now let $g(0) \neq 0$. Consider first the case where the genus of the canonical product belonging to $g(z)$ is 1. Then, since all zeros are negative and $g(x)$ is real,

$$g(z) = e^{Az} \prod_{m=1}^{\infty} (1 + r_m z) e^{-r_m z}$$

where

$$A \gtreqless 0, \; r_m > 0 \text{ and } \sum_{m=1}^{\infty} r_m^2 < +\infty.$$

Hence, on placing $z = b + it$, where $b > 0$ is fixed and $-\infty < t < +\infty$,

$$| g(b+it) |^2 = e^{2Ab} \prod_{m=1}^{\infty} [(1 + r_m b)^2 + r_m^2 t^2] e^{-2r_m b}.$$

This may be written in the form

$$(24) \qquad | g(b+it) |^{-2} = | g(b) |^{-2} \prod_{m=1}^{\infty} (1 + s_m^2 t^2)^{-1},$$

where

$$s_m = r_m/(1 + r_m b) > 0,$$

hence

$$\sum_{m=1}^{\infty} s_m^2 < +\infty, \text{ since } \sum_{m=1}^{\infty} r_m^2 < +\infty.$$

Thus it is clear that the arguments used in the Proof of Theorems I and II are applicable without change to the product (24). This proves Theorem VI for the case where the genus of the canonical product occurring in $g(z)$ is 1. If this genus is 0, then

$$g(z) = e^{Az} \prod_{m=1}^{\infty} (1 + r_m z),$$

where

$$A \gtreqless 0, \; r_m > 0 \text{ and } \sum_{m=1}^{\infty} r_m < +\infty.$$

Thus

$$| g(b + it) |^2 = e^{2Ab} \prod_{m=1}^{\infty} [(1 + r_m b)^2 + r_m^2 t^2].$$

Hence, on defining s_m by the same formula as in the previous case, it is easy to see that (24) is again valid. This completes the proof of Theorem VI.

The next Theorem will be an extension of Theorem I to the case of an "erhöht" genus, the latter being meant in the sense of Pólya [19].

THEOREM VII. *Let $F(z)$ be an entire function of the form*

$$F(z) = \exp(-bz^2) f(z),$$

where b is a real non-negative constant and $f(z)$ satisfies the general requirements of Theorem I but not necessarily the additional condition (14). *Then if $b > 0$, there exists a convex distribution function $\rho(x)$ such that*

$$F(0)/F(it) = L(t; \rho), \; -\infty < t < +\infty, \tag{25}$$

and ρ is a transcendental entire function which is bounded in every strip $-\alpha \leqq y \leqq \alpha$.

Proof. Theorem I belongs to $b = 0$. If $b > 0$, on replacing t by $b^{-1/2} t$ it may be assumed that $b = 1$. Thus

$$F(0)/F(it) = \exp(-t^2) f(0)/f(it) = \exp(-t^2) L(t; \sigma), \tag{26}$$

where σ is by Theorem I a convex distribution function. Since the even function $\exp(-x^2)$ is, up to constant factors, self-reciprocal under the Fourier cosine transform and decreases as $| x |$ increases, there exists a convex distribution function $\omega(x)$ such that $\exp(-t^2) = L(t; \omega)$, this distribution

being a symmetric Gaussian or normal distribution function. On placing $\rho = \omega * \sigma$, the distribution function ρ also is convex. Furthermore, ρ satisfies (25) in view of (26) and (4). Finally, since $|L(t;\sigma)| \leqq 1$ for every σ and for every t,

$$|L(t;\rho)| = |L(t;\omega)|\,|L(t;\sigma)| \leqq \exp(-t^2),$$

and so $L(t;\rho)$ satisfies (3) for a $\lambda > 1$. This completes the proof of Theorem VII.

The proof depended on the fact that if $\lambda = 2$, there exists a symmetric distribution function $\omega_\lambda = \omega_\lambda(x)$ such that

$$\exp(-|t|^\lambda) = L(t;\omega_\lambda), \quad -\infty < t < +\infty, \tag{27}$$

and that this $\omega = \omega_2$, being a symmetric Gaussian distribution function, is convex. According to Lévy [14], there exists a distribution function ω_λ satisfying (27) also when $0 < \lambda < 2$. If $\lambda = 1$, then one obtains the arcus-tangent distribution function, while if $0 < \lambda < 1$ or $1 < \lambda < 2$, then ω_λ is not an elementary function. All these distribution functions are symmetric, since $L(t;\omega_\lambda)$ is a real function. Now it will be shown that the distribution function ω_λ is convex not only in the trivial cases $\lambda = 1$ and $\lambda = 2$ but in the cases $0 < \lambda < 1$ and $1 < \lambda < 2$ also. The proof will require a modification of Lévy's generating distribution functions and will yield both the existence and the convexity of ω_λ.

THEOREM VIII. *The condition (27) defines for every positive $\lambda \leqq 2$ a convex distribution function ω_λ.*

Remark. Since there exists for every $\lambda > 0$ a function $\omega_\lambda(x)$ the derivative of which is

$$\omega'_\lambda(x) = \pi^{-1} \int_0^{+\infty} \exp(-t^\lambda) \cos xt\,dt, \tag{28}$$

it is clear from (1) and (2) that (27) may be satisfied by a real even function ω_λ also when $\lambda > 2$. If, however, $\lambda > 2$, then (28) attains (cf. Pólya [17], [18]) negative values for some real x; hence $\omega_\lambda(x)$ is not monotone non-decreasing and therefore not a distribution function. This fact will follow also from a more general result, to be proven later (Theorem XIV). Thus the point in Theorem VIII is that the distribution function ω_λ is convex whenever it exists. This fact will be the basis of § 3.

Information with regard to the analytic character of the distribution

functions ω_λ is given by the next Theorem. Theorems VIII and IX will be proven in § 5, where both appear as particular cases of more general results.

THEOREM IX. *The even function* $\omega_\lambda'(x)$ *defined by* (28) *for* $\lambda > 0$ *has for* $-\infty < x < +\infty$ *derivatives of arbitrarily high order and is regular at every real* $x \neq 0$. *If* $0 < \lambda < 1$, *there is a singularity at* $x = 0$, *which must be a transcendental singularity, since all derivatives exist at* $x = 0$ *also, if* x *is restricted to the real axis. Furthermore,* $x = 0$ *is a logarithmical branch point, if* λ *is sufficiently small. Thus if* $0 < \lambda < 1$, *the behavior of the function at* $x = 0$ *is about the same as that of Cauchy's standard example,* $\exp\{(\pm x)^\lambda\}$. *Correspondingly, it is not stated that the function elements* $\omega_\lambda'(x)$ *and* $\omega_\lambda'(-x)$, *which are regular for* $0 < x < +\infty$, *are analytic continuations of each other, although* $\omega_\lambda'(x)$ *is an even function. If, however,* $\lambda = 1$, *then* ω_λ' *is a rational function with two purely imaginary simple poles, while if* $\lambda > 1$, *then* ω_λ' *is a transcendental entire function of order* $(1 - \lambda^{-1})^{-1}$.

Remark. It is not known whether or not (3) implies, if $\lambda < 1$, the analyticity of $\sigma(x)$ at every inner point of the spectrum of $\sigma(x)$ at least in the simplest cases of infinite conventions; cf. Wintner [24], [25], Kershner and Wintner [13], and, in a multidimensional case, Jessen and Wintner [12], Theorems 14 and 19. Thus the analyticity of ω_λ for $x > 0$ is not obvious if $0 < \lambda < 1$.

The relation (27), when interpreted as an estimate of $L(t; \omega_\lambda)$ for large t, does not indicate what is the spectrum of ω_λ. For if (3) is satisfied, then the spectrum of σ need not be the whole line $-\infty < x < +\infty$ but may also be a bounded set. This holds for every fixed $\lambda < 1$, as shown by examples; cf. Wintner [25], [29]. Theorem IX implies, however,

COROLLARY I. *The spectrum of every distribution function* ω_λ *is the whole line* $-\infty < x < +\infty$.

For suppose the contrary. Then, since the spectrum always is a closed set, there exists an interval (a, b) such that $\omega_\lambda(x) = \text{const.}$, if $a < x < b$. Since $\omega_\lambda'(x)$ is regular both for $-\infty < x < 0$ and $0 < x < +\infty$ and since $\omega_\lambda'(x) = \omega_\lambda'(-x)$, it follows that $\omega_\lambda(x)$ is constant both for $-\infty < x < 0$ and $0 < x < +\infty$. Hence the spectrum either is empty or consists of the single point $x = 0$. The first case is impossible, since the total variation of every distribution function is 1. The second case is possible only for the unit distribution function χ, hence not for ω_λ. This contradiction proves Corollary I.

If $0 < \lambda \leqq 2$, then $\omega_\lambda(x)$ is a distribution function, hence $\omega_\lambda'(x) \geqq 0$,

and so (28) cannot have a real zero of odd multiplicity. It has been pointed out by Pólya ([18], p. 187, statement (a), where "finite" means "at most finite") that if $1 \leqq \lambda \leqq 2$, then (28) cannot have infinitely many real zeros. This result may be improved in two directions, as shown by

COROLLARY II. *If $0 < \lambda \leqq 2$, the function (28) has no real zeros.*

This fact, which is completed by the Remark which follows Theorem VIII, may be proven as follows. Suppose that there does exist a real $x = x_0$ such that $\omega_\lambda'(x_0) = 0$. Since $\omega_\lambda(x)$ is a distribution function, $\omega_\lambda'(x) \geqq 0$ for every x. Furthermore, if $|x'| < |x''|$, then $\omega_\lambda'(x') \geqq \omega_\lambda'(x'')$ in virtue of Theorem VIII. Consequently, $\omega_\lambda'(x) = 0$ for every x which is not in the finite interval $-|x_0| < x < |x_0|$. This consequence of the assumption $\omega_\lambda'(x_0) = 0$ contradicts Corollary I and proves therefore Corollary II.

Theorems VIII and IX will be proven in § 5, where the proofs are given for the case of more general integrals which depend on Bessel functions. The following remarks also concern Bessel functions but lie in another direction and are more elementary in character.

THEOREM IX. *There exist for every $\nu \geqq \frac{1}{2}$ two distribution functions, $\iota = \iota_\nu = \iota_\nu(x)$ and $\kappa = \kappa_\nu = \kappa_\nu(x)$, such that*

$$\text{const. } t^{-\nu} J_\nu(t) = L(t; \iota_\nu) \tag{29}$$

and

$$\text{Const. } t^{-\nu} \int_0^1 r^{\nu+1} J_\nu(tr)\, dr = L(t; \kappa_\nu), \tag{30}$$

where the factors of proportionality are positive and are determined by the condition $L(0; \cdot) = 1$. The spectrum is always the interval $-1 \leqq x \leqq 1$.

Remark. $\int_0^1$ may be replaced by $\int_0^R$, where $0 < R < +\infty$. In fact, this change replaces κ_ν by a distribution function which is similar to κ_ν in the sense of § 1.

Proof. Suppose first only $\nu > -\frac{1}{2}$ and put $\iota_\nu'(x)$ proportional to $(1 - x^2)^{\nu - 1/2}$, if $-1 < x < 1$, and $\iota_\nu'(x) = 0$, if $|x| \geqq 1$. Then (29) is, in view of the integral definition of the Bessel functions, satisfied by the symmetric distribution function ι_ν. Furthermore, ι_ν is a convex distribution function if and only if $\iota_\nu'(|x|)$ is monotone non-increasing, i. e., if and only if $\nu \geqq \frac{1}{2}$. It also follows that the left-hand side of (30) is

$$\int_0^1 \int_0^1 r^{2\nu+1}(1-s^2)^{\nu-1/2} \cos(trs)\,dr\,ds$$

up to a constant factor which is positive by the definition of $\iota_\nu'(x)$. On placing $rs = x$, where s is fixed, the last integral may be written in the form

$$\int_0^1 s^{-2\nu-2}(1-s^2)^{\nu-1/2}\{\int_0^s x^{2\nu+1}\cos(tx)\,dx\}ds.$$

On partial integration this becomes

$$\int_0^1 F_\nu(x)x^{2\nu+1}\cos(tx)\,dx,$$

i. e.,

$$\int_{-1}^1 \tfrac{1}{2}F_\nu(|x|)\,|x|^{2\nu+1}e^{itx}dx,$$

where

(31) $$F_\nu(x) = \int_x^1 s^{-2\nu-2}(1-s^2)^{\nu-1/2}ds;\ 0 < x < 1,$$

the infinity of $F_\nu(x)$ at $x = 0$ being compensated by the factor $x^{2\nu+1}$. Hence a distribution function $\kappa_\nu(x)$ satisfying (30) may be obtained by choosing $\kappa_\nu'(x)$ proportional to the positive function $F_\nu(|x|)\,|x|^{2\nu+1}$ or equal to zero according as $|x| < 1$ or $|x| \geqq 1$. Thus $\kappa_\nu(x)$ is a convex distribution function if and only if $F_\nu(x)x^{2\nu+1}$ is monotone non-increasing in the interval $0 \leqq x \leqq 1$. Hence it is sufficient to prove that

(32) $$\{F_\nu(x)x^{2\nu+1}\}' < 0, \text{ where } 0 < x < 1 \text{ and } ' = d/dx.$$

Now let $\nu \geqq \frac{1}{2}$. Then, from (31),

$$F_\nu(x) \leqq (1-x^2)^{\nu-1/2}\int_x^1 s^{-2\nu-2}ds,\ 0 < x < 1,$$

since $\nu - \frac{1}{2} \geqq 0$; and

$$(2\nu+1)\int_x^1 s^{-2\nu-2}ds = x^{-2\nu-1} - 1 < x^{-2\nu-1},$$

so that

(33) $$F_\nu(x) < (2\nu+1)^{-1}x^{-2\nu-1}(1-x^2)^{\nu-1/2},\ 0 < x < 1.$$

On calculating the derivative of the product $F_\nu(x)x^{2\nu+1}$ by using the definition (31) of $F_\nu(x)$, it is easily found that (33) may be written in the form (32). This completes the proof of Theorem X.

Since from (29)

(34) $$L(t;\iota_{1/2}) = 2^{1/2}\Gamma(\tfrac{1}{2}+1)t^{-1/2}J_{1/2}(t) = t^{-1}\sin t,$$

hence, by Vieta's product for $t^{-1}\sin t$,

$$L(t;\iota_{1/2})=\prod_{j=1}^{\infty}\cos(t/2^j),$$

it is easy to see that

$$\prod_{k=1}^{\infty}\cos(t/k)=L(2t;\iota_{1/2})L(2t/3;\iota_{1/2})L(2t/5;\iota_{1/2})\cdots.$$

Hence it is clear from the product rules (4) and (5) that the distribution function τ defined by (19) is

$$\iota_{1/2}(x/2)*\iota_{1/2}(3x/2)*\iota_{1/2}(5x/2)*\cdots, \tag{35}$$

and is therefore convex in view of the convexity of $\iota_{1/2}$. This suggests more general

Remarks on the convexity of the projections of spherical equidistributions. Let

$$b_1>0,\ b_2>0,\cdots \text{ and } b_1^2+b_2^2+\cdots<+\infty.$$

Then it is clear from (29) and (30) that both infinite products

$$\prod_{m=1}^{\infty}L(b_mt;\iota_\nu),\qquad \prod_{m=1}^{\infty}L(b_mt;\kappa_\nu)$$

are uniformly convergent in every fixed finite interval $|t|<\text{Const.}$ It therefore follows from the rule (5) that

$$\mathrm{I}_\nu(x)=\iota_\nu(x/b_1)*\iota_\nu(x/b_2)*\cdots \tag{36}$$

and

$$\mathrm{K}_\nu(x)=\kappa_\nu(x/b_1)*\kappa_\nu(x/b_2)*\cdots \tag{37}$$

are convergent infinite convolutions. Since the convexity is preserved by a similarly transformation, by the convolution process and by a limit process (cf. § 1), Theorem X implies that I_ν and K_ν are convex distribution functions for any $\{b_m\}$. This holds for every $\nu\geqq\frac{1}{2}$.

If $\nu=\frac{1}{2}n-1$, where n is an integer, the convexity of the distribution function (36) may be interpreted as a result on a random walk problem in an n-dimensional space, where $n\geqq 3$ in view of $\nu\geqq\frac{1}{2}$. In fact, the convexity of a distribution function means that the density of probability is a non-increasing function of the distance $|x|$ from the "median" $x=0$ (cf. § 3 below). Now $\iota_\nu(x/b)$ is the distribution function obtained by projection from the equidistribution on the sphere $\xi_1^2+\cdots+\xi_n^2=b^2$, where $n=2\nu+2$; cf.

Jessen and Wintner [12], § 3 and § 5, where further references are given. It is interesting that the infinite convolution (36) in the case of the dimension number $n = 2$, which belongs to $\nu = 0 < \frac{1}{2}$ and occurs in connection with almost periodic functions with linearly independent frequencies, is from the point of view of convexity just as exceptional as in the degenerate case $n = 1$; the latter belongs to $\nu = -\frac{1}{2}$, hence, in view of

$$(38) \qquad L(t; \iota_{-1/2}) = 2^{-1/2}\Gamma(-\tfrac{1}{2}+1)t^{1/2}J_{-1/2}(t) = \cos t,$$

to symmetric Bernoulli convolutions.

The convex distribution function K_ν defined by (37), where $\nu \geqq -\frac{1}{2}$, admits in the case $\nu = \frac{1}{2}n - 1$ of an interpretation similar to that of (36). In fact, $\kappa_\nu(x/b)$ is the distribution function obtained by projection from the equidistribution within the spherical volume $\xi_1^2 + \cdots + \xi_n^2 \leqq b^2$, where $n = 2\nu + 2$. For the Fourier transform L of the distribution function of the projection is easily found to be

$$\text{Const.} \int_0^b r^{n-1}(tr)^{1-n/2}J_{-1+n/2}(tr)\,dr,$$

an integral which is identical with $L(bt; \kappa_{-1+n/2})$ in view of (30). While it is clear from (38) and from the proof of Theorem X that if $n = 1$ and $n = 2$, then ι_ν is not convex, hence (36) need not be convex, the distribution function (37) is always convex in these cases also. In other words, $\kappa_{-1/2}$ and κ_0 are convex distribution functions. As far as $\nu = -\frac{1}{2}$, i. e., $n = 1$ is concerned, the truth of the statement is easily verified from (38), (34) and (30) and also is geometrically clear. If $n = 2$, then $\nu = 0 > -\frac{1}{2}$, and so the representation of $\kappa_\nu = \kappa_0$ in terms of (31) is valid. Hence it is sufficient to show that (32) holds for $\nu = 0$. Now from (31)

$$\{xF_0(x)\}' = \int_x^1 s^{-2}(1-s^2)^{-1/2}ds - x^{-1}(1-x^2)^{-1/2}.$$

Hence

$$\{xF_0(x)\}'' = -(1-x^2)^{-3/2} < 0; \quad 0 < x < 1.$$

Thus $\{xF_0(x)\}'$ is steadily decreasing and will, therefore, satisfy the condition (32), where $\nu = 0$, if it is negative in the neighborhood of $x = +0$. Now if $x \to +0$, then, by l'Hospital's rule,

$$\{xF_0(x)\}' \sim x\{xF_0(x)\}'' = -x(1-x^2)^{-3/2} < 0.$$

Thus the lowest dimension numbers, $n = 1$ and $n = 2$, are or are not exceptional with regard to convexity according as one considers the projection

of the equidistribution belonging to the boundary or to the interior of an n-dimensional sphere.

3. *On a postulate of Gauss.* Let Q' and Q'' be two places, Q a collinear place between them, d_0', d_0'', d_0 the "true" values and d', d'', d some measured values of the distances $Q'Q$, QQ'', $Q'Q''$ respectively. Instead of distances, one may think of masses or of arbitrary additive observables which are characterized by a real number. Suppose that the three "distances" have been measured very often. The result of a measurement will be, of course, almost never precisely the same as the result of a previous measurement, so that one has practically continuous series $d'-d_0'$, $d''-d_0''$, $d-d_0$ of errors of observation. If these series belong to reliable observations and have been obtained for practically every position of the intermediary point Q, the distribution of errors which are introduced by the limited accuracy of any measurement will satisfy three postulates of Gauss. The first of these conditions is satisfied for the normal law of Gauss (cf. Bessel [1]); Gauss [6] states the two others as general postulates. The three conditions in question will be conceded to be necessary for the case of "reliable" measurements, measurements where nothing is wrong with the instruments or with the observers.

(i) The three distribution curves plotted for the errors $d'-d_0'$, $d''-d_0''$ and $(d'+d'')-(d_0'+d_0'')$ must be similar to the distribution curve plotted for the error $d-d_0$ of the direct measurement of the distance $Q'Q''$, no matter what is the position of Q between Q' and Q''. The similarity of two distribution curves means, as in § 1, that they become identical by a suitable change of the unit of one of the two errors. As to this condition (i), cf. Bessel [1], Pólya [17].

(ii) The error $d-d_0$ must be just as probable and/or frequent as the error $-(d-d_0)$. In other words, the probability of an error must not depend on its sign but only on its magnitude.

(iii) If d_1 and d_2 are measured values of d_0 and $|d_1-d_0|<|d_2-d_0|$, the probability of the error d_1-d_0 must not be less than that of the error d_2-d_0. In other words, small errors must be at least as probable as large errors.

Conditions (ii) and (iii) together may be expressed by saying that the density of probability must be a non-increasing function of the absolute deviation from the "true" value. The latter was so far undefined but may now be defined as the median.

Postulates (i) and (ii) are independent of each other, since the asym-

metric arc tan-distributions of Cauchy [4] satisfy (i) without satisfying (ii). The question of whether all three postulates are independent of each other has apparently not been treated in the literature. In what follows, it will be shown that (iii) *is a consequence of* (i) *and* (ii). Conditions (i), (ii) and (iii) are, in the terminology of § 1, the conditions of stability, symmetry and convexity. Hence what one has to prove is

THEOREM XI. *Every symmetric stable distribution function is a convex distribution function.*

Remark. Since the symmetric distribution functions ω_λ defined by (27) satisfy the stability criterion (6), Theorem XI implies that every ω_λ is convex, as stated by Theorem VIII. On the other hand, the statement of Theorem XI, viz. that the postulates (i) and (ii) imply (iii), does not sound like an analytical statement but rather like one which, when true, must be provable without any analytical tools; hence the convexity of every ω_λ does not seem to be an analytical fact. While an approach to Theorem VIII from this abstract direction is perhaps not impossible, the way to be followed will be that of deducing Theorem XI from Theorem VIII. To this end it will be necessary to know that every symmetric stable distribution function is, in the main, an ω_λ. Cauchy [4], p. 101 and Lévy [14], p. 254, derived by summary considerations an exponential representation of the Fourier transform of an arbitrary stable distribution function, and this exponential representation shows that Theorem XI is implied by Theorem VIII. However, for a complete proof of Theorem XI a detailed proof of the exponential representation in question will be needed. This proof will also save a similar consideration in § 4. The proof of Theorem XI will occupy this whole § 3.

In order not to interrupt the proof in what follows, let it here be observed that if $l(s)$, where $0 < s < +\infty$, is a non-vanishing real or complex function such that for every $s > 0$, for every $v > 0$ and for some real constant $\lambda > 0$

$$(39) \qquad l(s^{1/\lambda}v) = [l(v)]^s, \text{ then } l(s) = \exp(-Cs^\lambda),$$

where C is a constant which need not be real. This statement is easily verified by choosing $v = 1$ and denoting by C one of the logarithms of $[l(1)]^{-1}$.

Now let σ be an arbitrary stable distribution function which is distinct from the unit distribution function χ, so that the positive number $c = c(a, b)$ defined at the end of § 1 is unique. Put $c_{m+1} = c(1, c_m)$, where $c_1 = 1$ and $m = 1, 2, \cdots$. Then

$$(40) \qquad [L(t;\sigma)]^m = L(c_m t;\sigma), \qquad (m = 1, 2, \cdots),$$

in view of (6). The positive sequence $\{c_m\}$ will be considered as undefined if $\sigma=\chi$, in which case (40) is satisfied by any c_m, since $L(t;\chi)=1$ for every t.

It is easy to prove that

$$(41)\qquad c_2>c_1=1 \text{ and } \limsup_{m=\infty} c_m=+\infty.$$

First, on choosing in (40) the exponent $m=2$ and iterating the resulting relation k times, it is seen that the 2^k-th power of $L(t;\sigma)$ is $L(c_2{}^k t;\sigma)$. Since there exists for every m but one c_m, it follows from (40) that $c_m=c_2{}^k$ where $m=2^k$ and $k=1,2,\cdots$. Hence if the first of the statements (41), viz. $c_2>1$, is true, then so is the second. Suppose, if possible, that $c_2>1$ is false. If $c_2=1$, then $L(t;\sigma)$ is, in view of (40), its own square for every t, since $c_1=1$. Hence $L(t;\sigma)$ is either 0 or 1 for a given t. Since $L(t;\sigma)$ is a continuous function, it follows that either $L(t;\sigma)=0$ for every t or $L(t;\sigma)=1$ for every t. The first case is impossible, since $L(0;\sigma)=1$ for every σ, and the second case is impossible, since $\sigma\neq\chi$ by supposition. Now let $c_2<1$. Then $L(c_2{}^k t;\sigma)$ tends, as $k\rightarrow+\infty$, to 1 for every t, since $L(0;\sigma)=1$. Hence the 2^k-th power of $L(t;\sigma)$ is 1 for every t, which again implies the contradiction $\sigma=\chi$. This completes the proof of (41).

It follows that

$$(42)\qquad L(t;\sigma)\neq 0,\ -\infty<t<+\infty,$$

for every stable σ. This is clear if $\sigma=\chi$. Let $\sigma\neq\chi$ and suppose, if possible, that $L(t_0;\sigma)=0$ for some $t=t_0$. Then it is clear from (40) that $L(t;\sigma)$ vanishes at $t=t_0/c_m$ also. Hence it is seen from the second of the relations (41) that $t=0$ is either a zero or a cluster point of zeros of $L(t;\sigma)$. Both cases are impossible, since $L(t;\sigma)$ is a continuous function and $L(0;\sigma)=1$. This proves (42).

Thus there exists for $-\infty<t<+\infty$ a unique continuous $\log L(t;\sigma)$ which vanishes at $t=0$. In what follows, $[L(t;\sigma)]^s$ will mean $\exp\{s\log L(t;\sigma)\}$ with the previous definition of the logarithm. Thus $[L(t;\sigma)]^s$ is a continuous function of t and s together.

Let m and n be arbitrary positive integers. Repeated application of (40) shows that

$$[L(tc_m/c_n;\sigma)]^n=L(tc_m;\sigma)=[L(t;\sigma)]^m.$$

Hence, on using the above definition of the power $[L(t;\sigma)]^s$,

$$(43)\qquad L(tc_m/c_n;\sigma)=[L(t;\sigma)]^{m/n}.$$

This implies that the Fourier transform of the distribution function $\sigma(ax)$,

where $a = c_m/c_n$, only depends on m/n. Since, as pointed out in § 1, the relation $\sigma(a_1 x) = \sigma(a_2 x)$ is impossible for $a_1 \neq a_2$ unless $\sigma = \chi$, and since $\sigma \neq \chi$ by assumption, it follows that c_n/c_m depends only on m/n.

Let s be an arbitrary positive number, $s_1 = n_1/m_1, \cdots, s_j = n_j/m_j, \cdots$ a sequence of rational positive numbers which tend to s as $j \to \infty$, and let $\sigma_j(x)$ denote the distribution function defined by

$$\sigma_j(x) = \sigma(x c_{m_j}/c_{n_j}),$$

so that

$$L(t; \sigma_j) = [L(t; \sigma)]^{s_j}$$

in view of (43). Thus, since $[L(t;\sigma)]^s$ is a continuous function of t and s together, it follows from the continuity theorem (§ 1) and from the assumption $s_j \to s$, where $j \to \infty$, that there exists a distribution function $\tau = \tau_s$ such that $\sigma_j \to \tau_s$ and

$$L(t; \tau_s) = [L(t; \sigma)]^s. \tag{44}$$

The last relation shows that $\tau_s(x)$ is uniquely determined by s, i. e. that $\tau_s(x)$ is independent of the sequence $\{s_j\}$ of rational numbers by which s has been approximated. Since $\sigma_j \to \tau_s$,

$$\sigma(x c_{m_j}/c_{n_j}) \to \tau_s(x), \quad j \to \infty,$$

by the definition of σ_j. Hence it is clear from $\sigma \neq \chi$ that c_{n_j}/c_{m_j} tends to a finite positive limit $f = f(s)$ for which

$$\sigma(x/f(s)) = \tau_s(x);$$

since $\tau_s(x)$ depends only on s, the same holds for the limit $f(s)$ of c_{n_j}/c_{m_j} in view of $\sigma \neq \chi$. Now the last representation of τ_s is equivalent to

$$L(f(s)t; \sigma) = L(t; \tau_s).$$

Consequently,

$$L(f(s)t; \sigma) = [L(t; \sigma)]^s \tag{45}$$

in view of (44). On using the continuity theorem (§ 1) and the assumption $\sigma \neq \chi$ once more, it is seen from (45) that the function $f(s)$, which is positive by its definition, is a continuous function of s, where $0 < s < +\infty$. Suppose that $f(s') = f(s'')$ for some pair of positive numbers s', s''. Then it is clear from (45) and (42) that either $s' - s'' = 0$ or $L(t; \sigma) = 1$ for every t. Since the second possibility is excluded by the assumption $\sigma \neq \chi$, it follows that $f(s') = f(s'')$ only when $s' = s''$. Consequently, $f(s)$ is a strictly monotone

function in view of its continuity. Since $c_m = f(m)$ in virtue of (40) and (45), it is clear from (41) that $f(s)$ is monotone increasing and $f(+\infty) = +\infty$.

The number c_m has so far been defined only for $m = 1, 2, \cdots$. If $s > 0$ is not an integer, put $c_s = c_m$, where m is the least integer exceeding s. Since $f(m/n) = c_m/c_n$ in view of (43) and (45); and since $f(s)$, where $0 < s < +\infty$, is a positive monotone continuous function, it is clear that if $s > 0$ is fixed and k is an integer which tends to $+\infty$, then $c_{sk}/c_k \to f(s)$, where sk is the product of s and k. Similarly, $c_{srk}/c_{rk} \to f(s)$, if $s > 0$ and $r > 0$ are fixed. Since c_{srk}/c_k is the product of c_{srk}/c_{rk} and c_{rk}/c_k, it follows by letting $k \to +\infty$ that $f(s)$ satisfies the functional equation $f(sr) = f(s)f(r)$. Hence, since $f(s)$ is positive and continuous, $f(s) = s^\kappa$, where κ is a real constant and $0 < s < +\infty$. Furthermore, $\kappa > 0$ in virtue of $f(+\infty) = +\infty$. On placing $\lambda = 1/\kappa$, it follows from (45) that

$$L(s^{1/\lambda}t;\sigma) = [L(t;\sigma)]^s, \qquad (s > 0,\ t \gtreqless 0), \tag{46}$$

where σ is an arbitrary stable distribution function and λ a positive constant depending on σ. If $\sigma = \chi$, then the relation (46) is but $1 = 1$, so that $\sigma = \chi$ need not be excluded. On choosing $t > 0$ and placing $t = v$, it is seen from (46) and (39) that there exists a constant C for which

$$L(t;\sigma) = \exp(-Ct^\lambda), \text{ if } 0 < t < +\infty. \tag{46a}$$

This holds for $t = 0$ also, since $L(0;\sigma) = 1$. Now $L(t;\sigma)$ and $L(-t;\sigma)$ are conjugated complex in view of (1). Hence

$$L(t;\sigma) = \exp(-\bar{C}\mid t\mid^\lambda), \text{ if } -\infty < t < 0. \tag{46b}$$

Now suppose that the stable distribution function σ is symmetric, i. e., that $\sigma(x) = 1 - \sigma(-x)$. Then $L(t;\sigma)$ is a real function in virtue of (1); hence $C = \bar{C}$, and so

$$L(t;\sigma) = \exp(-C\mid t\mid^\lambda),\ -\infty < t < +\infty,$$

where $\lambda > 0$. Furthermore, $C < 0$ is impossible, since $\mid L(t;\sigma)\mid \leqq 1$ for every σ and for every t. Thus either $C = 0$ or $C > 0$. If $C = 0$, then $L(t;\sigma) = 1$ for every t, hence $\sigma = \chi$. If $C > 0$, on replacing $\sigma(x)$ by $\sigma(C^{-1/\lambda}x)$ it may be assumed that $C = 1$. Hence if σ is symmetric and stable, then either $\sigma = \chi$ or σ is similar to a distribution function ω_λ which satisfies (27). Now if $\sigma(x)$ is a convex distribution function, then so is $\sigma(ax)$ for every $a > 0$. Furthermore, χ is a convex distribution function; cf. § 1. Hence Theorem XI is implied by Theorem VIII in view of the Remark which follows Theorem VIII.

4. *The multidimensional case.* Let $y = (y_1, \cdots, y_n)$ be a point of the real n-dimensional Cartesian space R_y. By a distribution function in R_y is meant an absolutely additive monotone set-function $\phi = \phi(E)$ which is defined for every Borel set E of R_y and is such that $\phi(R_y) = 1$. If the dimension number n of R_y is 1 and F_x denotes the Borel set $-\infty < y < x$, then $\sigma(x) = \phi(F_x)$ is a distribution function in the sense of § 1. Correspondingly, the Borel set E is termed in the case of an arbitrary n a continuity set of ϕ if $\phi(E_c) = \phi(E_i)$, where E_c denotes the closure of E and E_i the set of the interior points of E_i, hence possibly the empty set, in which case $\phi(E_i) = 0$. If ϕ is fixed, a set E is a continuity set of ϕ "in general"; namely in the same sense as a monotone function of a single variable is continuous "in general," that is to say up to a set of points which is at most enumerable. Correspondingly, a sequence $\{\phi_m\}$ of distribution functions in R_y is said to be convergent if there exists a distribution function ϕ such that $\phi_m(E) \to \phi(E)$ holds for every continuity set E of ϕ. The distribution function $\phi * \psi$ in R_y which represents the convolution of ϕ and ψ is defined in accordance with the case $n = 1$ of a single dimension (§ 1) and has the same properties as in the case $n = 1$. In particular, on placing

$$\Lambda(u; \phi) = \int_{Ry} e^{iuy} \phi(dR_y), \tag{47}$$

where the integral is a Radon integral, $u = (u_1, \cdots, u_n)$ is a point in a real space R_u and uy is the scalar product $u_1 y_1 + \cdots + u_n y_n$, one has

$$\Lambda(u; \phi * \psi) = \Lambda(u; \phi)\Lambda(u; \psi) \tag{48}$$

for every vector u in R_u. Furthermore, $\{\phi_m\}$ is a convergent sequence of distribution function if and only if the sequence $\{\Lambda(u; \phi_m)\}$ of the Fourier transforms is uniformly convergent within every fixed sphere of the space R_u, and $\lim \Lambda(u; \phi_m)$ is then $\Lambda(u; \lim \phi_m)$ for every u. This again will be referred to as the continuity theorem. A detailed treatment of the theory of the distribution functions in R_y is given by Haviland [9], [10], where further references may be found. The continuity theorem, when applied to the sequence $\phi, \psi, \phi, \psi, \phi, \cdots$, implies that there belongs to every Fourier transform $\Lambda(u; \phi)$ but one ϕ. For an explicit inversion formula cf. Haviland [10].

The distribution function ϕ is called absolutely continuous if there exists in R_y a measurable function $\delta(y) = \delta(y_1, \cdots, y_n)$ such that

$$\phi(E) = \int_E \delta(y)\mu(dR_y) \tag{49}$$

for every Borel set E, where $\mu(E)$ is the set-function r
Lebesgue measure in R_y, so that (49) is an ordinary
function $\delta(y)$, if it exists, is termed the density of $\phi($
on a set of measure zero, a remark which will not be
is monotone, $\delta(y) \geqq 0$. The inversion formula menti
if $|\Lambda(u;\phi)|$ has a finite integral over R_u, then ϕ is ab
a uniformly continuous and bounded density $\delta(y)$ whi
(47) for every y by the ordinary Fourier inversion,

$$(50) \qquad \delta(y) = (2\pi)^{-n} \int_{Ru} e^{-iyu} \Lambda(u;\phi)\mu(d$$

Let O denote the origin of R_y when this single
Borel set. Thus if E is an arbitrary Borel set,
$\phi(EO) = 0$ according as the point O is or is not in
only if the point O is a continuity set of ϕ.

Let $\phi_\Omega = \phi_\Omega(E)$ denote the distribution functio
fixed orthogonal matrix representing a rotation of R_y
the Borel set into which E is turned by this transfor
Since the scalar product uy is invariant under rotatio
that $\Lambda(u;\phi_\Omega) = \Lambda(\Omega^{-1}u;\phi)$. In particular, $\Lambda(u;\phi$
if and only if $\Lambda(u;\phi)$ is a function of the length $|u|$
is a function of $|u|$ alone if and only if $\phi_\Omega = \phi$ for e
condition is satisfied, ϕ is said to be of radial symmetr
that an absolutely continuous ϕ is of radial symmetry
$\delta(y)$ is a function of $|y|$ alone. If in addition $|\Lambda$
has a finite integral over R_u, then

$$\delta(y) = B_n \int_0^{+\infty} \left\{ \int_0^{\pi} \exp.(-i\,|y|\,|u|\cos\vartheta)\sin^{n-2}\vartheta \right.$$
$$\times \Lambda(|u|;\phi)n$$

as seen from (50) by introducing polar coördinates, ϑ
the vectors y and u. This tacitly assumes that $n > 1$.
the inner integral in terms of Bessel functions and plac

$$(51) \quad \delta(y) = \delta(|y|)$$
$$= A_n |y|^{1-n/2} \int_0^{+\infty} s^{n/2} J_{-1+n/2}(|y|s)\Lambda(s;$$

a formula which holds, in view of (38) and (2), for n
then radial symmetry means symmetry in the sense c
version of (51) is the Cauchy-Poisson formula for sp

Let $|y|$ denote the length of the vector y. If there exists in the interval $0 < r < +\infty$ a monotone non-increasing function $\eta(r)$ such that

$$\phi(E) = \phi(EO) + \int_E \eta(|y|)\mu(dR_y) \tag{52}$$

for every Borel set E, then $\phi(E)$ will be said to be convex. This definition is justified by the fact that if $n = 1$, then $\phi(E)$ is convex if and only if $\sigma(x) = \phi(F_x)$ is a convex distribution function in the sense of § 1, the Borel set F_x being the interval $-\infty < y < x$. It is clear that every convex $\phi(E)$ is of radial symmetry. Also

$$\phi(E_r) = \phi(O) + \alpha_n \int_0^r \eta(s) s^{n-1} ds, \tag{53}$$

where E_r denotes the sphere $|y| < r$ and α_n a positive constant which depends only on the dimension number n. It will always be supposed that $r > 0$. On comparing (52), (53) with (49) it is seen that a convex ϕ is absolutely continuous if and only if $\phi(O) = 0$, in which case its density $\delta(y)$ is the non-increasing function $\eta(|y|) \geqq 0$. Hence it is clear that if ϕ is convex, there exist convex $\phi_1, \phi_2, \cdots$ such that on the one hand $\phi_m \to \phi$ and on the other hand every ϕ_m has a density $\delta_m(y) = \delta_m(r)$ which possesses a continuous derivative in the interval $0 < r < +\infty$.

THEOREM XII. *If $\phi_1, \phi_2, \cdots$ are convex and $\phi_m \to \phi$, then ϕ also is convex.*

Proof. Since ϕ_m is convex, hence of radial symmetry, $\phi = \lim \phi_m$ clearly is of radial symmetry. Consequently, (52) is implied by (53), and so it is sufficient to prove that condition (53) is satisfied by a non-negative non-increasing η. Since ϕ_m is convex,

$$\phi_m(E_r) = \phi_m(O) + \alpha_n \int_0^r \eta_m(s) s^{n-1} ds, \tag{54}$$

where η_m is non-negative and non-increasing. Thus, since $0 \leqq \phi_m \leqq 1$,

$$1 \geqq \alpha_n \int_0^r \eta_m(s) s^{n-1} ds \geqq \alpha_n \int_\epsilon^r \eta_m(s) s^{n-1} ds,$$

where $\epsilon > 0$ is arbitrarily fixed and $\epsilon \leqq r < +\infty$. The integrand is the product of two monotone functions which increase, if $n > 1$, in opposite directions. It is, however, seen from the last inequality that

$$\alpha_n^{-1} \geqq \eta_m(2\epsilon) \int_\epsilon^{2\epsilon} s^{n-1} ds.$$

Hence the sequence $\{\eta_m(2\epsilon)\}$ is bounded, implying that $\{\eta_m(r)\}$ is uniformly bounded for $2\epsilon \leqq r < +\infty$, where $\epsilon > 0$ is arbitrarily small. Thus it follows from Helly's compactness theorem that the sequence $\{\eta_m(r)\}$ contains a subsequence which is convergent in the whole range $0 < r < +\infty$. The corresponding subsequence of $\{\eta_m(r)r^{n-1}\}$ is uniformly bounded in every finite interval $0 < a \leqq r \leqq b$ and may, therefore, be integrated term by term. Consequently, since $\phi_m \rightarrow \phi$, it is seen from (54) that

$$(55) \qquad \phi(E_b) - \phi(E_a) = \alpha_n \int_a^b \eta(s) s^{n-1} ds,$$

where η denotes the limit of the selected subsequence of $\{\eta_m\}$ and is, therefore, a non-negative and non-increasing function of $r > 0$. Since $a > 0$ and $b > 0$ are arbitrary, (55) clearly implies (53). This completes the proof of Theorem XII. Incidentally, a standard argument shows that the selection is, in reality, superfluous, since $\eta_m(r) \rightarrow \eta(r)$ at every continuity point r of $\eta(r)$. Of course, $\phi_m \rightarrow \phi$ does not imply $\phi_m(O) \rightarrow \phi(O)$, since $E = O$ need not be a continuity set of $\phi(E)$.

THEOREM XIII. *If* $\phi_1(E)$ *and* $\phi_2(E)$ *are convex, then so is* $\phi_1(E) * \phi_2(E)$.

Proof. It is clear from (48) and from the continuity theorem of the Fourier transform (47) that $\phi_{1m} \rightarrow \phi_1$ and $\phi_{2m} \rightarrow \phi_2$ imply $\phi_{1m} * \phi_{2m} \rightarrow \phi_1 * \phi_2$. Hence, by the remark which precedes Theorem XII, it is sufficient to prove Theorem XIII under the restriction that $\phi_i(E)$, where $i = 1$ and $i = 2$, is absolutely continuous and has a density $\delta_i(y) = \delta_i(|\,y\,|) = \delta_i(r)$ which admits of a continuous derivative for $0 < r < +\infty$. Since $\delta_i(r)$ is non-increasing, $\delta_i'(r) \leqq 0$, while, of course, $\delta_i(r) \geqq 0$. Since $\phi_i(E)$ is of radial symmetry, $\Lambda(u; \phi_i)$ depends only on $|\,u\,|$. Hence $\Lambda(u; \phi_1 * \phi_2)$, as product of $\Lambda(u; \phi_1)$ and $\Lambda(u; \phi_2)$, is a function of $|\,u\,|$ alone. Thus $\phi_1 * \phi_2$ is of radial symmetry. Since ϕ_i is absolutely continuous, so is $\phi_1 * \phi_2$ (cf., e. g., Kershner and Wintner [13], p. 543, footnote, where $n = 1$ but the proof holds for any n). Let $\delta(y) = \delta(|\,y\,|) = \delta(r)$ denote the density of $\phi_1 * \phi_2$. Then, by the definition of a convolution (cf. Haviland [9]),

$$\delta(|\,y\,|) = \delta(y) = \int_{Rv} \delta_1(y - v)\delta_2(v)\mu(dR_v),$$

where $y - v$ is the difference of the vectors y and v. Since the functions $\delta_i(r)$, $\delta_i'(r)$ are of constant sign, it is clear that $\delta'(r)$ exists and may be obtained by formal differentiation,

$$\delta'(r) = \delta'(|y|) = \int_{Rv} \delta_1'(y - v)\delta_2(v)\mu(dR_v),$$

the differentiation of $\delta_1(y - v) = \delta_1(|y - v|)$ being one with respect to the radius vector. Let H_v and $R_v - H_v$ be the two halves of R_v into which R_v is cut by a plane through the origin of R_v, and let H_v be chosen such that it contains the arbitrarily fixed point $y \neq (0, \cdots, 0)$. On writing $y - v$ instead of v as integration variable, it is seen from $\delta_i(y) = \delta_i(|y|) = \delta_i(-y)$ and from the directional character of the differentiation that

$$\delta'(r) = \delta'(|y|) = \int_{Hv} \delta_1'(|v|)\{\delta_2(|y - v|) - \delta_2(|y + v|)\}\mu(dR_v).$$

For if v is in H_v, then $-v$ is in $R_v - H_v$. Now y is in H_v. Hence y lies nearer to every point v of H_v than to the corresponding point $-v$ of $R_v - H_v$. Thus $|y - v| < |y + v|$ for every v in H_v. Since $\delta_2'(r) \leqq 0$, it follows that $\delta_2(|y - v|) \geqq \delta_2(|y + v|)$ for every v in H_v, i. e., that the difference $\{\ \}$ in the last integral is non-negative in the whole domain of integration. The factor $\delta_1'(|v|)$ of the integrand is nowhere positive. Hence $\delta'(r) \leqq 0$, where r is arbitrary. Consequently, $\delta(r)$ is monotone non-increasing. Since $\phi_1 * \phi_2$ is absolutely continuous, hence vanishes for $E = O$, the proof of Theorem XIII is herewith completed in view of (52), (53) and (49).

If $u \neq (0, \cdots, 0)$, put $e_u = u/|u|$, so that e_u is the unit vector which has the same direction as u. The next Theorem contains an extension of a fact mentioned in the Remark which follows Theorem VIII.

THEOREM XIV. *If a distribution function ϕ is such that, for some constant $\lambda > 0$ and for some function $C = C(e_u)$ of the direction,*

$$\Lambda(u; \phi) = \exp\{-C(e_u)|u|^\lambda\}, \tag{56}$$

then either $\lambda \leqq 2$ or $C(e_u) = 0$ for every u; in the latter case ϕ is the unit distribution function χ.

Proof. It is seen from (47) that $|\Lambda(u; \phi)| \leqq 1$ for every u and $\Lambda(u; \phi) = 1$ for $u = (0, \cdots, 0)$. Since

$$|\exp\{-C(e_u)\}| = |\Lambda(e_u; \phi)| \leqq 1,$$

the real part of $C(e_u)$ is bounded from below. Hence, on taking the real part of

$$\int_{Ry} (1 - e^{iuy})\phi(dR_y) = 1 - \Lambda(u; \phi),$$

it is clear from (56) that

$$\int_{R_y} (1 - \cos uy)\phi(dR_y) = O(|u|^\lambda) \text{ as } |u| \to 0.$$

Thus, since $\alpha^{-2}(1 - \cos\alpha)$ has in the interval $-1 \leqq \alpha \leqq 1$ a positive minimum, it follows from

$$0 \leqq \int_{|u|\,|y| \leqq 1} \leqq \int_{R_y} = O(|u|^\lambda)$$

that

$$(57) \qquad \int_{|u|\,|y| \leqq 1} |y|^2 \phi(dR_y) = O(|u|^{\lambda-2}) \text{ as } |u| \to 0,$$

the integration being extended over the sphere $|y| \leqq |u|^{-1}$ of R_y. This sphere tends to R_y as $|u| \to 0$. Now if $\lambda > 2$, then $O(|u|^{\lambda-2}) \to 0$ as $|u| \to 0$; hence, in view of (57),

$$\int_{R_y} |y|^2 \phi(dR_y) = 0,$$

which is possible only when the spectrum of ϕ consists of the origin of R_y, i. e., when $\phi = \chi$. This completes the proof of Theorem XIV.

The limiting case $\lambda = 2$ may be characterized by means of the "standard deviation" (dispersion)

$$(58) \qquad \mathrm{M}(\phi) = \int_{R_y} |y|^2 \phi(dR_y) = \int_{R_y} \sum_{j=1}^{n} |y_j|^2 \phi(dR_y) \leqq +\infty$$

as follows:

THEOREM XV. *A given distribution function ϕ satisfies* (56) *for $\lambda = 2$ and for some function $C(e_u)$ of the direction if and only if $0 \leqq \mathrm{M}(\phi) < +\infty$, where $\mathrm{M}(\phi) = 0$ only when $\phi = \chi$.*

Proof. If $\lambda = 2$, then $O(|u|^{\lambda-2})$ remains bounded as $|u| \to 0$, hence $\mathrm{M}(\phi) < +\infty$ in view of (57). Conversely, let $\mathrm{M}(\phi) < +\infty$. Then it is clear from (58) that the integrals

$$(59) \qquad \mu_j = \int_{R_y} y_j \phi(dR_y),\ \mu_{pq} = \int_{R_y} y_p y_q \phi(dR_y), \text{ where } j, p, q = 1, 2, \cdots, n,$$

are absolutely convergent in virtue of the Schwarz inequality. It follows therefore from (47) that $\Lambda(u; \phi)$ has at every point $u = (u_1, \cdots, u_n)$ of R_u continuous partial derivatives of the first and second order, and that these derivatives may be obtained from (47) by differentiation beneath the integral sign. Hence, on applying to $\Lambda(u; \phi)$ Taylor's formula in the neighborhood of $u = (0, \cdots, 0)$,

$$\Lambda(u;\phi)=1+i\sum_{j=1}^{n}\mu_j u_j-\tfrac{1}{2}\sum_{p=1}^{n}\sum_{q=1}^{n}\mu_{pq}u_p u_q+o\,(\sum_{j=1}^{n}u_j^2)$$

as $|u|\rightarrow 0$. On the other hand,

$$\Lambda(u;\phi)=1-C(e_u)\{(\sum_{j=1}^{n}u_j^2)^{\lambda/2}+o\,((\sum_{j=1}^{n}u_j^2)^{\lambda/2})\},\ |u|\rightarrow 0,$$

in view of (56). Now suppose that $\phi\neq\chi$. Then, since $\Lambda(u;\phi)=1$ for every u only when $\phi=\chi$, the factor $C(e_u)$ occurring in the second approximate representation of $\Lambda(u;\phi)$ does not vanish identically. Furthermore, it is clear from (58) and (59) that $\phi\neq\chi$ implies $\mathrm{M}(\phi)>0$, hence $\mu_{pp}(\phi)>0$ for at least one p, so that the quadratic form occurring in the first approximate representation of $\Lambda(u;\phi)$ does not vanish identically. Consequently,

$$\sum_{j=1}^{n}\mu_j u_j=0\quad\text{and}\quad-\tfrac{1}{2}\sum_{p=1}^{n}\sum_{q=1}^{n}\mu_{pq}u_p u_q=C(e_u)\,(\sum_{j=1}^{n}u_j^2)^{\lambda/2},$$

for every u. Since the last relation implies $\lambda=2$, Theorem XV follows.

It also follows that the matrix $\|-\tfrac{1}{2}\mu_{pq}\|$ is $C(e_u)$ times the unit matrix. Consequently, since the integrals (59) are independent of u, hence of e_u, the function $C(e_u)$ also is independent of e_u. This leads to

THEOREM XVI. *If a distribution function ϕ satisfies (56) for $\lambda=2$ and for some function $C=C(e_u)$ of the direction, then C is a real non-negative constant.*

Proof. As just shown, $C(e_u)$ is independent of e_u. Hence

$$\Lambda(u;\phi)=\exp(-C\,|u|^2),$$

where C is a constant. Since $|\Lambda(u;\phi)|\leqq 1$ in view of (47), it follows by letting $|u|\rightarrow\infty$ that the real part of C is non-negative. This implies Theorem XVI, since $\Lambda(u;\phi)$ and $\Lambda(-u;\phi)$ are conjugated complex in view of (47).

Remark. Since (56) implies

$$\Lambda(u;\phi)\sim 1-C(e_u)\,|u|^{\lambda},\ \text{where}\ |u|\rightarrow 0,$$

it is obvious that $\Lambda(u;\phi)$ has at the single point $u=(0,\cdot\cdot\cdot,0)$ partial derivatives of order $[\lambda]$, where $[\lambda]$ denotes the least integer not exceeding λ. This fact may be extended in the usual way to the case of fractional differentiation, only that the order must then be replaced by $\lambda-\epsilon$. On the other hand, it is *not* obvious that the derivatives at $u=(0,\cdot\cdot\cdot,0)$ may be obtained by *formal* differentiation of the integral (47). In fact, it is not clear that

$$\int_{Ry} | y |^{\nu} \phi(dR_y) < +\infty$$

for $\nu = \lambda$ or for every positive $\nu < \lambda$. Otherwise Theorem XV would hardly need a proof. As far as this difficulty is concerned, there is no real difference between Lévy's case $n = 1$ and the case of an arbitrary dimension number n.

Theorem XVI has been stated for $n = 1$ by Lévy [14], p. 261, footnote (2).

If c is a positive number, let $\phi_c(E)$ denote the distribution function $\phi(cE)$, it being understood that cE is the set of all points $cy = (cy_1, \cdots, cy_n)$, where y is an arbitrary point of E. It is clear from (47) that $\Lambda(u; \phi_c) = \Lambda(u/c; \phi)$. As in the case $n = 1$ of § 1, and for the same reason, $\phi(E)$ will be said to be a stable distribution function if there exists for every $a > 0$ and every $b > 0$ a $c = c(a, b) > 0$ such that $\phi_a * \phi_b = \phi_c$, i. e.,

$$\Lambda(u/a; \phi)\Lambda(u/b; \phi) = \Lambda(u/c; \phi). \tag{60}$$

THEOREM XVII. *A distribution function ϕ is stable if and only if $\Lambda(u; \phi)$ may be represented in the form* (56), *where the positive constant λ and the function $C(e_u)$ of the direction depend on ϕ.*

Remark. It is not stated, and it is not true, that if λ and $C(e_u)$ are given, (56) may be satisfied by a distribution function ϕ; Theorem XVII only decides when is a *given* ϕ a stable ϕ. For $n = 1$ cf. Lévy [14].

Proof. If (56) is satisfied, on placing

$$c = c(a, b) = (a^{-\lambda} + b^{-\lambda})^{-1/\lambda} > 0, \text{where } a > 0 \text{ and } b > 0,$$

(60) is clearly satisfied, and so ϕ is stable. Conversely, suppose that ϕ is stable. Then there exists a constant $\lambda > 0$ such that

$$\Lambda(s^{1/\lambda}u; \phi) = [\Lambda(u; \phi)]^s (\neq 0)$$

for every $s > 0$ and for every vector u. This is the extension of (46) for the case of an arbitrary dimension number n and is proven exactly in the same way as in the case $n = 1$ of § 3. Now, since $u = | u | e_u$,

$$\Lambda(s^{1/\lambda} | u | e_u; \phi) = [\Lambda(| u | e_u; \phi)]^s.$$

Hence on placing

$$v = | u | \text{ and } l(v) = \Lambda(ve_u; \phi),$$

where e_u is arbitrarily fixed, (56) follows from (39). This completes the proof of Theorem XVII.

The next Theorem might be of some physical interest in the case $n = 3$.

THEOREM XVIII. *If the dispersion* (58) *of a stable distribution function* ϕ *is neither zero nor infinite, then* ϕ *must be the Gauss-Maxwell law of radial symmetry.*

Proof. By Theorem XVII, $\Lambda(u;\phi)$ is of the form (56). Since $\mathrm{M}(\phi) < +\infty$ by assumption, $\lambda = 2$ and $C(e_u)$ is a real non-negative constant by Theorems XV and XVI. This constant cannot be zero, since otherwise $\Lambda(u;\phi) = 1$ for every u, i. e., $\phi = \chi$, which is excluded by the assumption $\mathrm{M}(\phi) > 0$. Thus $\Lambda(u;\phi) = \exp(-C \mid u \mid^2)$, where C is a positive constant. This implies Theorem XVIII, since $\exp(-r^2)$ is, up to constant factors, self-reciprocal under the Fourier cosine transformation.

The problem discussed at the beginning of § 3 may be extended to the multidimensional case. The postulate which corresponds to (iii) is again implied by the two other postulates, as shown by

THEOREM XIX. *Every stable distribution function of radial symmetry is convex.*

Proof. Let ϕ be stable. Then $\Lambda(u;\phi)$ is of the form (56) in view of Theorem XVII. Thus if ϕ is of radial symmetry, and so $\Lambda(u;\phi)$ a function of $\mid u \mid$ alone, then $C(e_u)$ is a constant. It is shown as in the Proof of Theorem XVI that this constant C is positive or zero. If $C = 1$, then $\Lambda(u;\phi) = \exp(-\mid u \mid^{\lambda})$, where $0 < \lambda \leqq 2$ in view of Theorem XIV, and it will be shown by Theorem XX that this ϕ exists and is convex for $n = 1, 2, \cdots$ and $0 < \lambda \leqq 2$. The case $C > 0$ may be reduced to the case $C = 1$ by a change of the length of the unit in R_y. This completes the proof of Theorem XIX, since if $C = 0$, then $\phi = \chi$, and χ is a convex distribution function.

5. *Cauchy's transcendents and their generalizations.* Theorems VIII and IX have not been proven in § 2. Theorems XX and XXI, to be proven in what follows, generalize Theorems VIII and IX to the case of an arbitrary dimension number.

THEOREM XX. *There exists for every dimension number* n *and for every positive* $\lambda \leqq 2$ *a convex distribution function* $\phi = \phi_\lambda = \phi_\lambda(E)$ *such that*

$$\Lambda(u;\phi_\lambda) = \exp(-\mid u \mid^{\lambda}). \tag{61}$$

Remark. If $\lambda > 2$, the distribution function ϕ_λ does not exist in virtue of Theorem XIV.

Proof. If $\lambda = 2$, the existence and the convexity of ϕ_λ is obvious; cf. the end of the Proof of Theorem XVIII. Let therefore $0 < \lambda < 2$. Since $\lambda > 0$, the integral of the positive function $\mathrm{Min}(1, \mid y \mid^{-n-\lambda})$ over the whole space R_y

is finite, and so there exists a distribution function $\psi_\lambda = \psi_\lambda(E)$ the density of which is

(62) $$A_{n\lambda}\,\mathrm{Min}(1, |\,y\,|^{-n-\lambda}), \text{ where } A_{n\lambda} > 0,$$

the factor $A_{n\lambda}$ of proportionality being determined by the condition $\psi_\lambda(R_y) = 1$. Let ϑ denote the angle between the vectors u and y, so that

$$uy = sr\cos\vartheta, \text{ where } s = |\,u\,|,\ r = |\,y\,|,\ 0 \leqq \vartheta \leqq \pi.$$

Since

$$\Lambda(u;\psi_\lambda) = A_{n\lambda}\{\int_{|y|\leqq 1} e^{iuy}\mu(dR_y) + \int_{1\leqq|y|} e^{iuy}\,|\,y\,|^{-n-\lambda}\mu(dR_y)\}$$

in view of (47), (49) and (62), on introducing into R_y polar coördinates and denoting by B_n a positive constant which depends only on the dimension number n, it is seen that

$$\Lambda(u;\psi_\lambda) = A_{n\lambda}\Big\{B_n\int_0^1\Big[\int_0^\pi \exp(i\,|\,u\,|\,r\cos\vartheta)\sin^{n-2}\vartheta\,d\vartheta\Big]\,nr^{n-1}dr$$
$$+ B_n\int_1^{+\infty}\Big[\int_0^\pi \exp(i\,|\,u\,|\,r\cos\vartheta)\sin^{n-2}\vartheta\,d\vartheta\Big]\,nr^{n-1}\cdot r^{-n-\lambda}dr\Big\}.$$

It will be convenient to use the abbreviation

(63) $$\{J_\nu(z)\} = 2^\nu\Gamma(\nu+1)z^{-\nu}J_\nu(z), \qquad \nu \geqq -\tfrac{1}{2},$$

so that $\{J_\nu(x)\}$ is an even entire function. Since both inner integrals represent $C_n\{J_{-1+n/2}(|\,u\,|\,r)\}$, where C_n is a positive constant, it follows by placing $\alpha_{n\lambda} = A_{n\lambda}B_n nC_n$ that

(64) $$\Lambda(u;\psi_\lambda)/\alpha_{n\lambda} = \int_0^1 r^{n-1}\{J_{-1+n/2}(|\,u\,|\,r)\}dr + \int_1^{+\infty}\{J_{-1+n/2}(|\,u\,|\,r)\}r^{-1-\lambda}dr,$$

where $\alpha_{n\lambda}$ is a positive constant. So far it has been tacitly assumed that $n > 1$, since if $n = 1$, there is no polar angle ϑ varying from $\vartheta = 0$ to $\vartheta = \pi$. It is, however, clear from (38) and (63) that (64) holds for $n = 1$ also. Since by (63) and by the integral definition of the Bessel function

(65) $$\{J_\nu(z)\} = \pi^{-1/2}(\Gamma(\nu+1)/\Gamma(\nu+\tfrac{1}{2}))\int_{-1}^{1} e^{iz\theta}(1-\theta^2)^{\nu-1/2}d\theta,$$

if $\nu > -\frac{1}{2}$, it is clear that

(65a) $$|\{J_\nu(x)\}| \leqq \{J_\nu(0)\} > 0;\ -\infty < x < +\infty,$$

and (65a) holds, in view of (38), for $\nu = -\frac{1}{2}$ also.

Put

(66) $$f_{\lambda n}(q) = \int_q^{+\infty} v^{-1-\lambda}[\{J_{-1+n/2}(0)\} - \{J_{-1+n/2}(v)\}]dv, \text{ where } q > 0.$$

Since $\{J_\nu(z)\}$ is an even entire function, the integrand in (66) is, as $v \to +0$,

$$v^{-1-\lambda}O(v^2) = O(v^{1-\lambda}) = O(v^{-1+\epsilon}),$$

where $\epsilon > 0$ in virtue of $\lambda < 2$. Thus there exists a finite limit $f_{\lambda n}(+0)$. Furthermore, this limit is positive. In fact, the integrand in (66) is almost everywhere positive in view of (65a), which means that $f_{\lambda n}(q)$ is a steadily decreasing positive function of $q > 0$, so that the limit $f_{\lambda n}(+0)$ also is positive. Thus, on placing $\beta_{n\lambda} = f_{\lambda n}(+0)$,

(67) $$|u|^\lambda f_{\lambda n}(|u|) = \beta_{n\lambda}|u|^\lambda + o(|u|^\lambda) \text{ as } |u| \to 0,$$

where $\beta_{n\lambda} > 0$. Similarly, if

(68) $$h_n(q) = q^{-n}\int_0^q v^{n-1}[\{J_{-1+n/2}(0)\} - \{J_{-1+n/2}(v)\}]dv, \text{ where } q > 0,$$

then, as $q \to +0$,

$$h_n(q) = q^{-n}\int_0^q v^{n-1}\,O(v^2)dv = O(q^2),$$

and so, since $\lambda < 2$,

(69) $$h_n(|u|) = o(|u|^\lambda) \text{ as } |u| \to 0.$$

On choosing in (64) the point u as the origin of R_u and combining the resulting relation with (64) itself, it is seen that

(70) $$1/\alpha_{n\lambda} - \Lambda(u;\psi_\lambda)/\alpha_{n\lambda}$$

is the sum of

(71) $$\int_0^1 r^{n-1}\{J_{-1+n/2}(0)\}dr - \int_0^1 r^{n-1}\{J_{-1+n/2}(|u|r)\}dr$$

and

(72) $$\int_1^{+\infty}\{J_{-1+n/2}(0)\}r^{-1-\lambda}dr - \int_1^{+\infty}\{J_{-1+n/2}(|u|r)\}r^{-1-\lambda}dr.$$

On placing $v = |u|r$, where $|u| \neq 0$ is fixed, (71) and (72) appear in the respective forms

$$|u|^{-n}\int_0^{|u|} v^{n-1}[\{J_{-1+n/2}(0)\} - \{J_{-1+n/2}(v)\}]dv = h_n(|u|)$$

and

$$|u|^\lambda\int_{|u|}^{+\infty}[\{J_{-1+n/2}(0)\} - \{J_{-1+n/2}(v)\}]v^{-1-\lambda}dv = |u|^\lambda f_{\lambda n}(|u|)$$

in view of (68) and (66). Hence the difference (70) is

$$= h_n(|u|) + |u|^\lambda f_{\lambda n}(|u|),$$

and so, according to (69) and (67),

$$= o(|u|^\lambda) + \beta_{n\lambda}|u|^\lambda + o(|u|^\lambda) \text{ as } |u| \to 0.$$

Consequently, on denoting by $\gamma = \gamma_{n\lambda}$ the product of the positive constants $\alpha_{n\lambda}$ and $\beta_{n\lambda}$ and placing $a = \gamma^{1/\lambda} > 0$,

$$(73) \qquad \Lambda(u;\psi_\lambda) = 1 - a^\lambda|u|^\lambda + o(|u|^\lambda) \text{ as } |u| \to 0; \quad a > 0.$$

It follows that ψ_λ may be applied, in the manner of Lie, as an infinitesimal generator of the distribution function ϕ_λ the existence and convexity of which are to be proven. For let $\psi_{m\lambda}$ denote the distribution function defined by

$$(74) \qquad \psi_{m\lambda}(E) = \psi_\lambda(am^{1/\lambda}E) * \psi_\lambda(am^{1/\lambda}E) * \cdots * \psi_\lambda(am^{1/\lambda}E),$$

where the "factor" ψ_λ occurs m times. Then from (48)

$$\Lambda(u;\psi_{m\lambda}) = [\Lambda(a^{-1}m^{-1/\lambda}u;\psi_\lambda)]^m,$$

since the Fourier transform of $\phi(cE)$ is $\Lambda(u/c;\phi)$. Consequently, from (73),

$$\Lambda(u;\psi_{m\lambda}) = [1 - m^{-1}|u|^\lambda + o(m^{-1}|u|^\lambda)]^m, \quad |u| \to 0.$$

Hence it is clear from a standard property of the exponential function that

$$\Lambda(u;\psi_{m\lambda}) \to \exp(-|u|^\lambda), \quad m \to +\infty,$$

holds uniformly in every fixed sphere $|u| <$ Const. of R_u. It therefore follows from the continuity theorem mentioned at the beginning of § 4 that there exists a distribution function $\phi_\lambda = \lim \psi_{m\lambda}$ for which $\Lambda(u;\phi_\lambda)$ is the limit of $\Lambda(u;\psi_{m\lambda})$ as $m \to +\infty$. Thus ϕ_λ satisfies (61). Furthermore, $\phi_\lambda = \lim \psi_{m\lambda}$ is, according to Theorem XII, certainly convex if every $\psi_{m\lambda}$ is convex, while (74) is, according to Theorem XIII, certainly convex if $\psi_\lambda(cE)$, where $c = a^{-1}m^{1/\lambda}$, is convex, i. e., if $\psi_\lambda(E)$ is convex. Now $\psi_\lambda(E)$ has been defined as the absolutely continuous distribution function the density of which is (62), and (62) clearly is a non-increasing function of $|y|$; hence $\psi_\lambda(E)$ is convex (cf. § 4). This completes the proof of Theorem XX.

Since the integral of (61) over the whole space R_u is finite, (51) is applicable. Hence the density of $\phi_\lambda(E)$ is $A_n F_{n\lambda}(|y|)$, where $A_n > 0$ and

$$(75) \qquad F_{n\lambda}(x) = x^{1-n/2}\int_0^{+\infty} s^{n/2}\exp(-s^\lambda)J_{-1+n/2}(xs)\,ds = F_{n\lambda}(-x),$$

since $\{J_\nu(z)\}$ is an even function. If $n = 1$, this density goes over into (28) in virtue of (38). Hence Theorem IX is implied by

THEOREM XXI. *The even function* (75), *where* $n = 1, 2, \cdots$, *has for* $-\infty < x < +\infty$ *derivatives of arbitrarily high order and is regular at every real* $x \neq 0$. *If* $0 < \lambda < 1$, *there is at* $x = 0$ *a singularity which has the*

character of the one described in Theorem IX. If $\lambda = 1$, the function $F_{n\lambda}(z) = F_{n\lambda}(x+iy)$ is regular in a strip $|y| <$ const. but not in the whole plane, while if $\lambda > 1$, it is a transcendental entire function of order $(1-\lambda^{-1})^{-1}$. Finally, the function has no real zero at all or has real zeros of odd multiplicity according as $0 < \lambda \leqq 2$ or $\lambda > 2$.

Remark. The proof of the last statement is, in view of Theorems XX and XIV, exactly the same as that given in the case $n = 1$ in § 2. Except for the last statement of Theorem XXI, it will be needless to assume that n is an integer. On placing $\nu = -1 + \frac{1}{2}n$ and using the abbreviation (63), the function (75) is

$$(76) \qquad G_{\nu\lambda}(z) = \int_0^{+\infty} \{J_\nu(zs)\} s^{2\nu+1} \exp(-s^\lambda)\, ds$$

up to a constant positive factor. In what follows, instead of $\nu = -\frac{1}{2}, 0, \frac{1}{2}, 1, \cdots$, merely $\nu \geqq -\frac{1}{2}$ will be asssumed, while, as before, $\lambda > 0$.

Proof. It is clear from (65a) that the integral

$$(77) \qquad H_{\nu\lambda}(z) = \int_0^{+\infty} \{J_\nu(s)\} s^{2\nu+1} \exp(-s^\lambda/z^\lambda)\, ds$$

is, for every fixed real or complex $z \neq 0$, majorized by the integral

$$(78) \qquad \{J_\nu(0)\} \int_0^{+\infty} s^{2\nu+1} \,|\exp(-s^\lambda/z^\lambda)|\, ds.$$

Let W_λ denote the infinite open wedge

$$(79) \qquad W_\lambda:\quad -\tfrac{1}{2}\pi/\lambda < \operatorname{arc} z < \tfrac{1}{2}\pi/\lambda, \text{ where } z \neq 0.$$

Thus W_λ is simply connected but not necessarily schlicht; in fact, the number of its sheets becomes infinite as $\lambda \rightarrow 0$. If $\lambda < 1$, the half-plane consisting of all numbers $z \neq 0$ which have a non-negative real part lies within W_λ. Let z^λ be defined in W_λ as the univalued and continuous function of the position which is positive along the half-line $\operatorname{arc} z = 0$. Then the integral (78) is uniformly convergent in every fixed closed and bounded sub-set of (79). Hence, since (77) is majorized by (78) and z^λ is regular on W_λ, the integral (77) represents a regular function $H_{\nu\lambda}(z)$ on W_λ. As far as the integral (76) is concerned, one cannot say more than that it is uniformly convergent along the real axis, a property which does not indicate the analyticity of the function $G_{\nu\lambda}(z)$. If $0 < \lambda < 1$, it is clear from

$$\log \operatorname*{Max}_{|z| \leqq r} |\{J_\nu(z)\}| = \log \{J_\nu(ir)\} \sim r; \quad r \rightarrow \infty,$$

that the integral (76) is divergent at all points $z \neq 0$ of the imaginary axis $\arc z = \pm \frac{1}{2}\pi$, points which are contained in W_λ if $\lambda < 1$. However,

(80) $$G_{\nu\lambda}(z) = z^{-2\nu-2} H_{\nu\lambda}(z)$$

for every $z > 0$, as seen from (76) and (77) by a change of the integration variable. Since $H_{\nu\lambda}(z)$ is a regular function on W_λ, the relation (80) gives a regular analytic continuation of $G_{\nu\lambda}(z)$, where $0 < z < +\infty$, into W_λ. This holds for every $\lambda > 0$. Since $\{J_\nu(z)\} = \{J_\nu(-z)\}$, the function (76) also is even, and so it is not necessary to consider its analytic continuation which belongs to $-\infty < z < 0$.

It is clear from (65) and (38) that the m-th derivative of $\{J_\nu(x)\}$ is bounded for $-\infty < x < +\infty$, where m is arbitrarily fixed. Hence the integral resulting from (76) by differentiating the integrand m times with respect to z is uniformly convergent for all real z. Thus if z is real, all derivatives of the function (76) exist and may be obtained by formal differentiation; in particular,

(81) $$G_{\nu\lambda}{}^{(m)}(0) = \{J_\nu(0)\}^{(m)} \int_0^{+\infty} s^m s^{2\nu+1} \exp(-s^\lambda)\, ds$$

for every m, it being understood that these derivatives are obtained by restricting z to the real axis. On combining this with the analytic continuation obtained before, it follows that Theorem XXI will be proven if one shows that the radius of convergence of the power series

(82) $$\sum_{m=0}^{\infty} m!^{-1} G_{\nu\lambda}{}^{(m)}(0) z^m$$

is zero, finite and positive or infinite according as $0 < \lambda < 1$, $\lambda = 1$ or $\lambda > 1$, and that the order is $(1-\lambda^{-1})^{-1}$ in the latter case. In fact, if $\lambda \geqq 1$, then the integral (76) clearly is uniformly convergent at least in a small circle $|z| <$ const., so that the series (82) necessarily represents the function $G_{\nu\lambda}(z)$ within the circle of convergence.

Now if $\nu > -\frac{1}{2}$, then from (65)

$$\{J_\nu(0)\}^{(m)} = \pi^{-1/2}(\Gamma(\nu+1)/\Gamma(\nu+\tfrac{1}{2}))i^m \int_{-1}^{1} \theta^m (1-\theta^2)^{\nu-1/2} d\theta;$$

hence

(83) $$\{J_\nu(0)\}^{(2m+1)} = 0,$$

while

$$\{J_\nu(0)\}^{(2m)} = (-1)^m \pi^{-1/2} \Gamma(\nu+1)\Gamma(m+\tfrac{1}{2})/\Gamma(m+\nu+1),$$

as seen by introducing into $\int_{-1}^{1} = 2\int_0^1$ the integration variable θ^2. It is

clear from (38) and (63) that (83) holds for $\nu = -\frac{1}{2}$ also. On the other hand, the factor of $\{J_\nu(0)\}^{(m)}$ in (81) is

$$\int_0^{+\infty} s^{(m+2\nu+1)/\lambda} \exp(-s)\lambda^{-1}s^{-1+1/\lambda}\, ds = \lambda^{-1}\Gamma(2\lambda^{-1}(\tfrac{1}{2}m+\nu+1)).$$

Hence

$$(84) \quad G_{\nu\lambda}^{(2m)}(0) = (-1)^m g_{\nu\lambda}\Gamma(2\lambda^{-1}(m+\nu+1))\Gamma(m+\tfrac{1}{2})/\Gamma(m+\nu+1),$$

where

$$g_{\nu\lambda} = \pi^{-1/2}\lambda^{-1}\Gamma(\nu+1)$$

is positive and independent of m; furthermore, $G_{\nu\lambda}^{(2m+1)}(0) = 0$. Now it is clear from (84) and from Stirling's formula that

$$|G_{\nu\lambda}^{(2m)}(0)|^{1/(2m)} \sim [\Gamma(2m/\lambda)]^{1/(2m)} \text{ as } m \to \infty.$$

Consequently, the reciprocal value of the radius of convergence of (82) is

$$\limsup_{m=\infty} |G_{\nu\lambda}^{(m)}(0)/m!|^{1/m} = \lim_{m=\infty} [\Gamma(2m/\lambda)/\Gamma(2m)]^{1/(2m)}$$

and this is, by Stirling's formula, 0, 1, or $+\infty$ according as $\lambda > 1$, $\lambda = 1$ or $\lambda < 1$. Finally, it is known that the order of an entire function $f(z)$ may be obtained merely from the sequence of its Taylor coefficients by calculating

$$\limsup_{m=\infty} (m \log m)/\log|m!/f^{(m)}(0)|.$$

This completes the proof of Theorem XXI, since $G_{\nu\lambda}^{(2m+1)}(0) = 0$, while

$$\log|(2m)!/G_{\nu\lambda}^{(2m)}(0)| \sim \log[\Gamma(2m)/\Gamma(2m/\lambda)] \sim (1-\lambda^{-1})[2m\log(2m)],$$

as seen from (84) and from Stirling's formula.

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] F. W. Bessel, *Abhandlungen*, vol. 2 (1876), pp. 383-384.

[2] E. W. Cannon and A. Wintner, "An asymptotic formula for a class of distribution functions," *Proceedings of the Edinburgh Mathematical Society*, ser. 2, vol. 4 (1935), pp. 138-143.

[3] T. Carleman, *Les fonctions quasi analytiques*, Paris, 1926.

[4] A. Cauchy, *Oeuvres complètes*, ser. 1, vol. 12 (1900), pp. 94-114, more particularly pp. 99-101.

[5] E. C. Francis and J. E. Littlewood, *Examples in infinite series with solutions*, Cambridge, 1928.
[6] C. F. Gauss, *Werke*, vol. 4 (1873), p. 5, § 4 and p. 7, § 6.
[7] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge, 1934.
[8] F. Hausdorff, "Beitraege zur Wahrscheinlichkeitsrechnung," *Berichte ueber die Verhandlungen der Koenigl. Saechsischen Gesellschaft der Wissenschaften zu Leipzig*, vol. 53 (1901), pp. 152-178.
[9] E. K. Haviland, "On the theory of absolutely additive distribution functions," *American Journal of Mathematics*, vol. 56 (1934), pp. 625-658.
[10] ———, "On the inversion formula for Fourier-Stieltjes transforms in more than one dimension," *ibid.*, vol. 57 (1935), pp. 94-100 and pp. 382-388.
[11] E. K. Haviland and A. Wintner, "On the Fouirer-Stieltjes transform," *ibid.*, vol. 56 (1934), pp. 1-7.
[12] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.
[13] R. Kershner and A. Wintner, "On symmetric Bernoulli convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 541-548.
[14] P. Lévy, *Calcul des probabilités*, Paris, 1925, more particularly Chapter VI of the second part.
[15] M. Mathias, "Ueber positive Fourier-Integrale," *Mathematische Zeitschrift*, vol. 16 (1923), pp. 103-125; cf. also G. A. Campbell and R. M. Foster, *Fourier Integrals for Practical Applications*, 1931, p. 41, pair 301. 1.
[16] G. Pólya, "Ueber die Nullstellen gewisser ganzen Funktionen," *Mathematische Zeitschrift*, vol. 2 (1918), pp. 352-383.
[17] ———, "Herleitung des Gaussschen Fehlergesetzes aus einer Funktionalgleichung," *Mathematische Zeitschrift*, vol. 18 (1923), pp. 96-108.
[18] ———, "On the zeros of an integral function represented by Fourier's integral," *Messenger of Mathematics*, vol. 52 (1923), pp. 185-188.
[19] ———, "Ueber die algebraisch-funktionentheoritischen Untersuchungen von J. L. W. V. Jensen," *Kgl. Danske Videnskabernes Selskab, Meddelelser*, vol. 7, no. 17 (1927).
[20] E. C. Titchmarsh, *The zeta-function of Riemann*, Cambridge, 1930.
[21] N. Wiener, *The Fourier integral and certain of its applications*, Cambridge, 1933.
[22] J. R. Wilton, "Note on the zeros of Riemann's zeta-function," *Messenger of Mathematics*, vol. 45 (1916), pp. 180-183.
[23] A. Wintner, "On the stable distribution laws," *American Journal of Mathematics*, vol. 55 (1933), pp. 335-339.
[24] ———, "Upon a statistical method in the theory of diophantine approximations," *ibid.*, vol. 55 (1933), pp. 309-331.
[25] ———, "On analytic convolutions of Bernoulli distributions," *ibid.*, vol. 56 (1934), pp. 659-663.
[26] ———, "On symmetric Bernoulli convolutions," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 137-138.
[27] ———, "A note on the Riemann chsi-function," *The Journal of the London Mathematical Society*, vol. 10 (1935), pp. 82-83.
[28] ———, "Gaussian distributions and convergent infinite convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 821-826.
[29] ———, "On convergent Poisson convolutions," *ibid.*, vol. 57 (1935), pp. 827-838.
[30] A. Zygmund, *Trigonometrical series*, Warszawa and Lwów, 1935.

ON THE ASYMPTOTIC DISTRIBUTION OF ALMOST PERIODIC FUNCTIONS WITH LINEARLY INDEPENDENT FREQUENCIES.

By Richard Kershner and Aurel Wintner.

Let

$$x(t) \sim \sum_{m=0}^{\infty} a_m \cos \lambda_m (t - \delta_m), \text{ where } a_m \geqq 0 \text{ and } 0 \leqq \lambda_m \delta_m < 2\pi,$$

be almost periodic in the sense of Bohr and let $A \leqq x \leqq B$ denote the least closed interval containing all values x attained by $x(t)$ for $-\infty < t < +\infty$. It has been proven [1] that $x(t)$ always posseses an asymptotic distribution function $\sigma(x)$, $-\infty < x < +\infty$, and that this monotone function $\sigma(x)$ is nowhere constant in the interval $A \leqq x \leqq B$, while, of course, $\sigma(x) \equiv 0$ if $x < A$ and $\sigma(x) \equiv 1$ if $x > B$. The present note deals with the important particular case where $a_m \neq 0$ for every m and the frequencies $\lambda_0, \lambda_1, \cdots$ are linearly independent, so that

$$-A = B = \sum_{m=0}^{\infty} a_m < +\infty \qquad (a_m > 0)$$

by a well-known theorem of Bohr.

On denoting by $\rho(x)$, $-\infty < x < +\infty$, the monotone continuous function which is 0, 1 or $1 - \pi^{-1} \operatorname{arc\,cos} x$ according as x is on the left, on the right or in the interior of the interval $-1 < x < 1$ and placing, for $-\infty < x < +\infty$,

$$\sigma_{n+1}(x) = \int_{-\infty}^{+\infty} \sigma_n(x - \xi)\, d\rho_{n+1}(\xi), \tag{1}$$

$$\text{where } \rho_n(x) = \rho(x/a_n) \text{ and } \sigma_0(x) = \rho_0(x),$$

the asymptotic distribution function $\sigma(x)$ of $x(t)$ may be obtained from [2]

$$\sigma(x) = \lim_{n=+\infty} \sigma_n(x) \tag{2}$$

and has, for $-\infty < x < +\infty$, derivatives of arbitrarily high order.[2] Hence there arises the question of whether $\sigma(x)$ is analytic.[3] There is a general method [4] by which the analyticity of a distribution function $\sigma(x)$ can be proven in some cases but this method breaks down in the present case. For the

---

[1] A. Wintner, *Spektraltheorie der unendlichen Matrizen*, Leipzig, 1929, pp. 254-255 and p. 269.

[2] A. Wintner, *American Journal of Mathematics*, vol. 55 (1933), pp. 312-316.

[3] Cf. *ibid.*, p. 310.

[4] A. Wintner, *American Journal of Mathematics*, vol. 56 (1934), p. 659.

method in question yields the analyticity of $\sigma(x)$ either for every x or for no x. Now the present $\sigma(x)$ cannot be analytic at the end points of the interval

$$(3) \qquad -\sum_{m=0}^{\infty} a_m \leqq x \leqq \sum_{m=0}^{\infty} a_m.$$

In fact, $\sigma(x)$ is nowhere constant in this interval but is constant both on the left and on the right of it, so that all derivatives of $\sigma(x)$ vanish at both end points of (3). The object of the present note is *to delimit in the range* (3) *of* $x(t)$ *intervals within which the asymptotic distribution function* $\sigma(x)$ *is regular analytic.*

First, $\sigma(x)$ *is regular in the interval*

$$(4) \qquad -a_0 + \sum_{m=1}^{\infty} a_m < x < a_0 - \sum_{m=1}^{\infty} a_m$$

whenever $a_0 - \sum_{m=1}^{\infty} a_m > 0$.

In order to prove this, put

$$(5) \qquad q_{n+1} = q_n + a_{n+1}, \text{ where } q_0 = 0,$$

so that

$$(6) \qquad 0 < q_n < q_{n+1} \to q, \text{ where } q = \sum_{m=1}^{\infty} a_m.$$

Thus $a_0 - q > 0$ by assumption. Let ϵ be an arbitrarily small fixed positive number less than $a_0 - q$. On removing from a complex z-plane the half-lines $-\infty < z \leqq -1$ and $1 \leqq z < +\infty$, there results a simply connected schlicht domain in which the function $F(z) = (1 - z^2)^{-\frac{1}{2}}$, where $F(0) = +1$, is univalued and regular. If z is in the circle $|z| < a_0 - \epsilon$, the function $\sigma_0(z)$ defined by

$$\sigma_0(z) = (\pi a_0)^{-1} \int_{-a_0}^{z} F(\zeta/a_0)\, d\zeta$$

is regular and $|\sigma_0(z)|$ has in this circle a finite least upper bound M, since $\epsilon > 0$ is fixed. It is clear from (6) and from the assumption $\epsilon < a_0 - q$ that $a_0 - q_n - \epsilon > 0$, i. e., that $|z| < a_0 - q_n - \epsilon$ is a circle for every n. Suppose that, for a given $n = \bar{n}$, there has been defined in the circle $|z| < a_0 - q_n - \epsilon$ a function $\sigma_n(z)$ in such a way that $\sigma_n(z)$ is regular and $|\sigma_n(z)| \leqq M$ in this circle. This condition has been satisfied for $n = 0$, since $q_0 = 0$ by (5). If $n = \bar{n}$ is arbitrarily given and if $|z| < a_0 - q_{n+1} - \epsilon$ and $-a_{n+1} < \xi < a_{n+1}$, then $|z - \xi| < a_{\bar{n}} - q_n - \epsilon$ in view of (5). Hence one may define in the circle $|z| < a - q_{n+1} - \epsilon$ a function $\sigma_{n+1}(z)$ by placing

$$\sigma_{n+1}(z) = (\pi a_{n+1})^{-1} \int_{-a_{n+1}}^{a_{n+1}} \sigma_n(z - \xi) F(\xi/a_{n+1})\, d\xi.$$

Then, since $\sigma_n(z)$ is regular and $|\sigma_n(z)| \leqq M$ in the circle $|z| < a - q_n - \epsilon$, it is clear that $\sigma_{n+1}(z)$ is regular and

$$|\sigma_{n+1}(z)| \leqq (\pi a_{n+1})^{-1} \int_{-a_{n+1}}^{a_{n+1}} M \, |F(\xi/a_{n+1})| \, d\xi = M\pi^{-1} \int_{-1}^{1} (1-\xi^2)^{-\frac{1}{2}} \, d\xi = M$$

in the circle $|z| < a_0 - q_{n+1} - \epsilon$.

Thus there has been defined a sequence $\{\sigma_n(z)\}$ such that $\sigma_n(z)$ is regular and $|\sigma_n(z)| \leqq M$ in the circle $|z| < a_0 - q_n - \epsilon$, where $n = 0, 1, \cdots$. Hence it is seen from (6) that the functions $\sigma_n(z)$ are regular and uniformly bounded in the circle $|z| < a_0 - q - \epsilon$ which is independent of n. Now it is clear from the successive definition of the complex functions $\sigma_n(z)$ that, when z is real, they are identical with the functions defined by (1), so that $\{\sigma_n(z)\}$ is convergent for real z in view of (2). It follows, therefore, from the elements of the theory of normal families (Vitali) that $\{\sigma_n(z)\}$ is uniformly convergent in every closed subset of the circle $|z| < a_0 - q - \epsilon$. Since ϵ is arbitrarily small, the limit function is regular in the circle $|z| < a_0 - q$. In particular, (2) is regular in the interval $-a_0 + q < x < a_0 - q$ or, by the definition (6) of q, in the interval (4); q. e. d.

The assumption of the result thus proven was that the amplitudes a_m of $x(t)$ satisfy the condition

$$(7) \qquad a_n > \sum_{m=n+1}^{\infty} a_m$$

for $n = 0$. If (7) holds not only for $n = 0$ but for $n = 1$ also, an iteration [5] of the above argument shows that $\sigma(x)$ is regular analytic not only in the interval (4) but also in each of the two additional subintervals

$$\pm a_0 - a_1 + \sum_{m=2}^{\infty} a_m < x < \pm a_0 + a_1 - \sum_{m=2}^{\infty} a_m$$

of (3) which have no point in common with each other or with (4). It suffices to notice that the singular points of $\sigma_1(x)$ are the four points $\pm a_0 \pm a_1$, the signs being this time independent of each other. Similarly, if (7) holds for $n = 0, 1, \cdots, N$, one obtains $2^{N+1} - 1$ disjoint open subintervals of (3) such that (2) is regular analytic within each of these $2^{N+1} - 1$ intervals.

In order to describe the situation in the case $N = +\infty$ where (7) holds for every n, let the interval (3) be decomposed into two complementary sets S, T by placing a point x of (3) into S or into T according as x can or cannot be represented in at least one way in the form

[5] Cf. R. Kershner and A. Wintner, *American Journal of Mathematics*, vol. 57 (1935), pp. 544-545.

$$x = \sum_{m=0}^{\infty} \pm a_m,$$

where the signs are independent of each other. Then S is[6] perfect and nowhere dense, and so T is open and everywhere dense in the interval (3). Now it is easily seen that the intervals within which the indefinite iteration of the above argument yields the analyticity of the function (2) are exactly the intervals which constitute the open set T. *Thus if the amplitudes a_m of $x(t)$ satisfy the condition* (7) *for every n, then the asymptotic distribution $\sigma(x)$ is regular analytic in infinitely many disjoint open intervals which lie dense in the range* (3) *of $x(t)$*. It remains undecided whether or not the complementary set S, consisting of the cluster points of the end points of these open intervals of regularity, actually contains but singular points of $\sigma(x)$. All that is certain is that all derivatives of $\sigma(x)$ exist at every point of S also and that the set of the 2^{n+1} singular points $x = \pm a_0 \pm \cdots \pm a_n$ of $\sigma_n(x)$ tends to the set S as $n \to +\infty$, finally that the clustering of these singular points is strongest at the two end points of the interval (3), at which points $\sigma(x)$ certainly is not regular analytic (cf. the introduction). It may be mentioned[7] that the measure of the nowhere dense perfect set S is zero or positive according as $2^n a_n$ does or does not tend to zero as $n \to \infty$. If, for instance, $a_n = a^n$, where $0 < a < \frac{1}{2}$, then (7) holds for every n and S is of measure zero. This example is of interest in view of the non-differentiable function of Weierstrass.[8] If $a_n = 2^{-n} + 3^{-n}$, then (7) is satisfied for every n and S has a positive measure.

So far it has been assumed that $x(t)$ is real-valued. The case

$$x(t) + iy(t) \sim \sum_{n=0}^{\infty} a_n \exp i\lambda_n(t - \delta_n), \quad -\infty < t < +\infty,$$

of a complex-valued function offers no new problem, since it may be reduced to the case $y(t) \equiv 0$ treated above by means of an integral equation of the Abel type,[9] the frequencies λ_n being linearly independent. It is expected that the unsolved problems of analyticity in the general case of linearly independent moduli[10] may also be treated along the lines of the present note, the complications involved being but technical in nature.

THE JOHNS HOPKINS UNIVERSITY.

[6] Cf., e. g., *loc. cit.*[5]

[7] Cf. *loc. cit.*[5]

[8] Cf. A. Wintner, *American Journal of Mathematics*, vol. 55 (1933), pp. 603-605.

a positive measure.

[9] Cf. *loc. cit.*[2], pp. 316-319.

[10] Cf. *loc. cit.*[2], pp. 328-329 and, for a complete theory, B. Jessen and A. Wintner, *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

NECESSARY AND SUFFICIENT CONDITIONS FOR POTENTIALS OF SINGLE AND DOUBLE LAYERS.[1]

By George A. Garrett.

1. *Introduction.* We consider the problem of obtaining necessary and sufficient conditions that a harmonic function defined in a three-dimensional region T be such that it can be represented in the region as a potential function. Denoting the harmonic function by $v(M)$ or $U(M)$, we find conditions necessary and sufficient in order that the function be representable as the potential of a single-layer distribution,

$$v(M) = \int_{S_0} \frac{1}{MP}\, d\nu(e_P)$$

and as the potential of a double-layer distribution,

$$U(M) = \int_{S_0} \frac{\cos \sphericalangle\, (MP, n_P)}{\overline{MP}^2}\, d\nu(e_P);$$

where S_0 is the boundary of T; $\nu(e)$ is the mass function, a completely additive function of point sets defined on S_0; P is a point of S_0; and n_P is the interior normal to S_0 at P.

These conditions are given in terms of "normal families" of surfaces.

2. *Potentials and normal families of surfaces.*

Definition. We shall call a family $\{S\}$ of simple closed surfaces having a tangent plane at every point a normal family provided that they lie interior or exterior to the surface S_0, which we shall consider as a member of the family, and

(a) for each surface of the family and for any two points A, B of the surface

$$| \sphericalangle\, (n_A, n_B) | < \gamma \overline{AB}$$

where γ is a constant, n_A is the normal at A, and n_B is the normal at B;

[1] This paper is an extension of the corresponding treatment for the two-dimensional case as presented by Professor G. C. Evans in a seminar conducted at The Rice Institute during the spring of 1933. Theorems 4.1 and 4.2 overlap somewhat a theorem published by Ch. J. De la Vallée Poussin in November of the same year. Cf. De la Vallée Poussin, "Propriétés des Fonctions Harmoniques dans un Domaine Ouvert Limité par des Surfaces à Courbure Bornée," *Annali della Rendiconti Scuola Normale Superiore de Pisa,* serie 2, vol. 2 (1933-XI), pp. 167-199. A more complete reference to this theorem of De la Vallée Poussin's is given in the footnote to Theorem 4.2.

(b) for every two surfaces S_1 and S_2 of the family

$$|\sphericalangle (n_{A_1}, n_{A_2})| < \gamma \overline{A_1 A_2}$$

where n_{A_1} is the normal to S_1 at A_1 on S_1, and n_{A_2} is the normal to S_2 at A_2 on S_2;

(c) in any infinite subset of the family there is a subsequence $\{S'_i\}$ which approaches a surface S of the family, approach being in the sense that the maximum normal distance from S approaches zero.

We take the positive directions of the normals to the surfaces to be toward the interiors of the surfaces. We denote by T^+ and T^- the regions interior and exterior respectively to S_0.

Consider a surface S_0 having the property (a). Let P be a point of S_0 and M a point not on S_0 and such that the normal to S_0 at the point Q on S_0 passes through M. Let M' be the image of M in the tangent plane to S_0 at Q. Consider a normal family $\{S\}$ of surfaces inside or outside S_0 and let τ be the maximum normal distance of a member of the family from S_0. We suppose that τ is small enough so that there is a unique $1:1$ correspondence [2] of points A_0 of S_0 along the normals n_{A_0} with points A of an arbitrary surface of $\{S\}$. If M is on S, then we denote by n_M the normal to S at M. We suppose also that τ is so small that dS_Q/dS_M is bounded away from zero and infinity, where dS_Q and dS_M are the elements of area of S_0 at Q and S at M, respectively.

We write

$$\begin{aligned} \phi &= \sphericalangle (MP, n_P); & \phi' &= \sphericalangle (M'P, n_P) \\ \theta_M &= \sphericalangle (MP, n_M); & \theta'_M &= \sphericalangle (M'P, n_{M'}) \\ \theta_Q &= \sphericalangle (MP, n_Q); & \theta'_Q &= \sphericalangle (M'P, n_Q) \\ r &= MP; & r' &= M'P. \end{aligned}$$

By the symbol θ without subscripts we shall mean both θ_M and θ_Q; and by θ', both θ'_M and θ'_Q. We consider the functions

$$\text{(I).} \qquad U(M) = \int_{S_0} \frac{\cos\phi}{r^2}\, d\nu(e_P),$$

$$\text{(II).} \qquad U_1(M) = \int_{S_0} \frac{\cos\theta}{r^2}\, d\nu(e_P),$$

$$\text{(III).} \qquad U_2(M) = \int_{S_0} \frac{\sin\phi}{r^2}\, d\nu(e_P),$$

$$\text{(IV).} \qquad U_3(M) = \int_{S_0} \frac{\sin\theta}{r^2}\, d\nu(e_P).$$

---

[2] That this is true for τ small enough is shown by G. C. Evans and E. R. C. Miles in "Potentials of general masses in single and double layers. The relative boundary value problems," *American Journal of Mathematics*, vol. 53 (1931), pp. 493-516. Cf. p. 497. Hereafter we refer to this article as (A).

Let n_Q be the z-axis and the tangent plane to S_0 at Q be the xy-plane for a system of rectangular coördinates. Let s be the curve of intersection of S_0 and the plane determined by P and n_Q. We shall also denote by s the length of the arc QP measured along this curve. Take as the x-axis the tangent to s at Q. Let P have cylindrical coördinates (ρ, μ, z). Let A be any point of S_0. Then we denote by $S_0(\delta, A)$ the portion of S_0 containing A which is contained in the sphere of center A and radius δ; and by $S(\delta, A)$ the portion of S nearest to $S_0(\delta, A)$ which is cut out by this sphere. We suppose that P is in $S_0(\delta, Q)$ and that δ is so small that

$$s < 2\rho, \qquad \phi, \phi' \leq \pi.$$

In the inequalities which follow we may suppose without loss of generality that $r \leq r'$. For P in $S_0(\delta, Q)$ we have

$$
(2.1)\quad \left\{
\begin{aligned}
& QM = QM' \leq 2r; \\
& |\phi - \theta_Q| \leq \sphericalangle\,(n_P, n_Q); \quad |\phi' - \theta'_Q| \leq \sphericalangle\,(n_P, n_Q); \\
& |\theta_M - \theta_Q| \leq \sphericalangle\,(n_M, n_Q); \quad |\theta'_M - \theta'_Q| \leq \sphericalangle\,(n_{M'}, n_Q); \\
& |\cos\phi - \cos\theta_Q| \leq 2\left|\sin\frac{\phi - \theta_Q}{2}\right| < 2\gamma\rho; \\
& |\cos\phi' - \cos\theta'_Q| \leq 2\left|\sin\frac{\phi' - \theta'_Q}{2}\right| < 2\gamma\rho; \\
& |\sin\phi - \sin\theta_Q| \leq 2\left|\sin\frac{\phi - \theta_Q}{2}\right| < 2\gamma\rho; \\
& |\sin\phi' - \sin\theta'_Q| \leq 2\left|\sin\frac{\phi' - \theta'_Q}{2}\right| < 2\gamma\rho; \\
& |\sin\theta_Q - \sin\theta'_Q| \leq 2\left|\cos\frac{\theta_Q + \theta'_Q}{2}\right| \leq 2\,|\sin \sphericalangle\,(\rho, QP)| \\
& \qquad\qquad < \frac{2\gamma}{QP}\int_0^s s\,ds < 4\gamma\rho; \\
& |\cos\theta_Q + \cos\theta'_Q| \leq 2\left|\cos\frac{\theta_Q + \theta'_Q}{2}\right| < 4\gamma\rho; \\
& \left|\frac{1}{r^2} - \frac{1}{(r')^2}\right| \leq \left|\frac{\sin\theta_Q - \sin\theta'_Q}{\rho}\right| \left|\frac{1}{r} + \frac{1}{r'}\right| < 4\gamma\left(\frac{2}{r}\right) = \frac{8\gamma}{r}; \\
& |\sin\theta_M - \sin\theta_Q| \leq 2\left|\sin\frac{\theta_M - \theta_Q}{2}\right| < \gamma\,\overline{QM} \leq 2\gamma r; \\
& |\cos\theta_M - \cos\theta_Q| \leq 2\left|\sin\frac{\theta_M - \theta_Q}{2}\right| < 2\gamma r; \\
& |\sin\theta'_M - \sin\theta'_Q| \leq 2\left|\sin\frac{\theta'_M - \theta'_Q}{2}\right| < \gamma\,\overline{QM'} \leq 2\gamma r; \\
& |\cos\theta'_M - \cos\theta'_Q| \leq 2\left|\sin\frac{\theta'_M - \theta'_Q}{2}\right| < 2\gamma r.
\end{aligned}
\right.
$$

Since P is in $S_0(\delta, Q)$ if Q is in $S_0(\delta, P)$ the relations (2.1) subsist for Q in $S_0(\delta, P)$.

Now we wish to consider P as fixed and Q as a variable point in $S_0(\delta, P)$. Let n_P be the z'-axis and the tangent plane to S_0 at P be the $x'y'$-plane. Take the x'-axis along the tangent to the curve s^* at P, where s^* is the curve of intersection of S_0 and the plane determined by n_P and Q. We also denote by s^* the length of the arc QP measured along this curve. Let Q have cylindrical coördinates (ρ', μ', z'). Let $C(\delta, P)$ be the circle $\rho' = \delta$ in the $x'y'$-plane. Let δ be small enough so that the relations (2.1) hold and also small enough so that

$$s^* < 2\rho' \qquad |\, ds^* \,| < 2 \,|\, d\rho' \,|.$$

Hence we have

$$|\, dS_Q \,| < 2\rho' d\rho' d\mu' \qquad \tfrac{1}{2} < \rho/\rho' < 2.$$

LEMMA 1. *The integrals*

$$\int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta}{r^2} + \frac{\cos\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\sin\phi}{r^2} - \frac{\sin\phi'}{(r')^2} \right| dS_Q;$$

$$\int\limits_{S_0(\delta,P)} \left| \frac{\cos\phi - \cos\theta}{r^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\cos\phi' - \cos\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_M - \cos\theta_Q}{r^2} \right| dS_Q;$$

$$\int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta'_M - \cos\theta'_Q}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\sin\theta_M - \sin\theta_Q}{r^2} \right| dS_Q; \quad \int\limits_{S_0(\delta,P)} \left| \frac{\sin\theta'_M - \sin\theta'_Q}{(r')^2} \right| dS_Q;$$

all approach zero with δ, independently of τ. This statement is true if dS_Q and $S_0(\delta, P)$ are replaced by dS_M and $S(\delta, P)$ respectively.

Since dS_Q/dS_M is bounded away from zero and infinity, the second statement is a consequence of the first. Using θ_Q and θ'_Q for θ and θ' we have for the first integral

$$\begin{aligned} \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_Q}{r^2} + \frac{\cos\theta'_Q}{(r')^2} \right| dS_Q &= \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_Q + \cos\theta'_Q}{r^2} - \cos\theta'_Q \left(\frac{1}{r^2} - \frac{1}{(r')^2} \right) \right| dS_Q \\ &\leq \int\limits_{S_0(\delta,P)} \left\{ \left| \frac{\cos\theta_Q + \cos\theta'_Q}{r^2} \right| + \left| \frac{1}{r^2} - \frac{1}{(r')^2} \right| \right\} dS_Q \\ &< 12\gamma \int\limits_{S_0(\delta,P)} \frac{dS_Q}{\rho} \leq 48\gamma \int\limits_{C(\delta,P)} d\rho' d\mu' = 96\pi\gamma\delta. \end{aligned}$$

This inequality proves the lemma for the first integral. The four succeeding integrals are treated similarly, using θ_Q and θ'_Q for θ and θ'. For the sixth integral we have

$$\int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_M - \cos\theta_Q}{r^2} \right| dS_Q < 2\gamma \int\limits_{S_0(\delta,P)} \frac{dS_Q}{\rho} \leq 16\pi\gamma\delta.$$

The three remaining integrals are treated similarly. For the first integral using θ_M and θ'_M for θ and θ', we write

$$\int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_M}{r^2} + \frac{\cos\theta'_M}{(r')^2} \right| dS_Q \leq \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_M - \cos\theta_Q}{r^2} \right| dS_Q$$
$$+ \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta'_M - \cos\theta'_Q}{(r')^2} \right| dS_Q \quad + \int\limits_{S_0(\delta,P)} \left| \frac{\cos\theta_Q}{r^2} + \frac{\cos\theta'_Q}{(r')^2} \right| dS_Q.$$

The lemma has been proved for these last three integrals. We treat similarly the four succeeding integrals, using θ_M and θ'_M for θ and θ'.

As a corollary to Lemma 1, we have

LEMMA 2. *The integrals*

$$\int\limits_{S_0} \left| \frac{\cos\theta}{r^2} + \frac{\cos\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\sin\phi}{r^2} - \frac{\sin\phi'}{(r')^2} \right| dS_Q;$$
$$\int\limits_{S_0} \left| \frac{\cos\phi - \cos\theta}{r^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\cos\phi' - \cos\theta'}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\cos\theta_M - \cos\theta_Q}{r^2} \right| dS_Q;$$
$$\int\limits_{S_0} \left| \frac{\cos\theta'_M - \cos\phi'_Q}{(r')^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\sin\theta_M - \sin\theta_Q}{r^2} \right| dS_Q; \quad \int\limits_{S_0} \left| \frac{\sin\theta'_M - \sin\theta'_Q}{(r')^2} \right| dS_Q;$$

are bounded independently of τ. This statement is true if dS_Q and S_0 are replaced by dS_M and S respectively.

We denote by E_S a finite number of non-overlapping regions on S and let the corresponding regions on S_0 be E_{S_0}. We denote the measures of these sets by $m(E_S)$ and $m(E_{S_0})$.

THEOREM 2.1. *The absolute continuity of the integrals*

$$\int_{E_{S_0}} | U_3(M) - U_3(M') |\, dS_Q; \qquad \int_{E_{S_0}} | U_2(M) - U_2(M') |\, dS_Q;$$
$$\int_{E_S} | U_3(M) - U_3(M') |\, dS_M; \qquad \int_{E_S} | U_2(M) - U_2(M') |\, dS_M;$$

is uniform independently of τ.

It suffices to prove the theorem for the first two integrals, since dS_Q/dS_M is bounded away from zero and infinity. We write

$$I = \int_{E_{S_0}} | U_3(M) - U_3(M') | \, dS_Q \leq \int_{S_0} | d\nu(e_P) | \int_{E_{S}}$$

$$\leq \int_{S_0} | d\nu(e_P) | \left\{ \int_{S_0(\delta,P)} \left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| dS_Q + \int_{E_{S_0}-E_{S_0}\cdot} \left| \frac{\sin}{r} \right. \right.$$

We denote by $\bar{\nu}(S_0)$ the total variation of $\nu(e)$ over
pletely additive, $\bar{\nu}(S_0)$ is bounded. By Lemma 1, giv
δ so small that the first of the inner integrals is less th
constant K exists such that for Q in $E_{S_0} - E_{S_0} \cdot S_0($

$$\left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| < K.$$

We now have

$$I < \epsilon/2 + Km(E_{S_0})\bar{\nu}(S_0).$$

Hence, for all E_{S_0} such that $m(E_{S_0}) \leq \epsilon/2K\bar{\nu}(S_0)$, I
what was to be shown. The second integral may be

THEOREM 2. 2. *The integrals*

$$\int_{S_0} | U_3(M) - U_3(M') | \, dS_Q; \qquad \int_{S_0} | U_2(M)$$

$$\int_{S} | U_3(M) - U_3(M') | \, dS_M; \qquad \int_{S} | U_2(M)$$

remain bounded as τ approaches zero and approach ze

Again it suffices to prove the theorem for the first
the same method of proof is used for both. That the i
as τ approaches zero follows immediately from the p
the first integral we have

$$\int_{S_0} | U_3(M) - U_3(M') | \, dS_Q$$

$$\leq \int_{S_0} | d\nu(e_P) | \left\{ \int_{S_0 - S_0(\delta,P)} \left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| dS_Q + \int_{S_0(\delta,} \right|$$

By Lemma 1 we may fix δ so small that, given any $\epsilon >$

$$\int_{S_0(\delta,P)} \left| \frac{\sin\theta}{r^2} - \frac{\sin\theta'}{(r')^2} \right| dS_Q < \frac{\epsilon}{2\bar{\nu}(S}$$

independently of τ. Now for Q not in $S_0(\delta, P)$ the qua

is bounded for each τ. Let $K_1(\tau)$ be the least upper bound of this quantity. We have then

$$\int_{S_0} |U_3(M) - U_3(M')|\, dS_Q < K_1(\tau) m(S_0) \bar{v}(S_0) + \epsilon/2$$

for each τ. But $K_1(\tau)$ obviously approaches zero uniformly with τ. Hence the theorem is proved.

THEOREM 2. 3. *The integrals*

$$\int_S |U(M)|\, dS_M;\ \int_S |U_1(M)|\, dS_M;\ \int_{S_0} |U(M)|\, dS_Q;\ \int_{S_0} |U_1(M)|\, dS_Q$$

remain bounded as τ approaches zero. The same holds if $U(M')$, $U_1(M')$, are substituted for $U(M)$, $U_1(M)$, respectively.

It is sufficient to prove the theorem for the first two integrals. In fact, we need only to prove the theorem for the second integral, for

$$\int_S \big|\, |U(M)| - |U_1(M)|\, \big|\, dS_M \leq \int_S |U(M) - U_1(M)|\, dS_M$$
$$\leq \int_{S_0} |dv(e_P)| \int_S \left| \frac{\cos\phi - \cos\theta}{r^2} \right| dS_M$$

which is bounded by Lemma 2. Consider then

$$\int_S |U_1(M)|\, dS_M \leq \int_S dS_M \int_{S_0} \left| \frac{\cos\theta}{r^2} \right| dv(e_P).$$

Supposing for the moment that $\theta = \theta_M$, we have

$$\int_S \left| \frac{\cos\theta_M}{r^2} \right| dS_M = \int_S \left| \frac{\cos \sphericalangle (MP, n_M)}{r^2} \right| dS_M.$$

But S is a surface of "class Γ." [3] Hence a positive constant Γ exists so that $\int_S \left| \frac{\cos \sphericalangle (MP, n_M)}{r^2} \right| dS_M < \Gamma$. Hence we may interchange the order of integration above, obtaining

$$\int_S |U_1(M)|\, dS_M \leq \int_{S_0} |dv(e_P)| \int_S \frac{|\cos\theta_M|}{r^2}\, dS_M < \Gamma \bar{v}(S_0).$$

Therefore $\int_S |U_1(M)|\, dS_M$ is bounded for $\theta = \theta_M$. Moreover, for the pur-

[3] (A), p. 494.

poses of this theorem it is immaterial whether we use θ_M or θ_Q for θ, since $\int_S \left| \frac{\cos\theta_M - \cos\theta_Q}{r^2} \right| dS_M$ is bounded, by Lemma 2.

For the second part of the theorem, we may write

$$| U(M') | \leq | U(M') + U(M) | + | U(M) | ;$$
$$| U_1(M) | \leq | U(M') | + | U_1(M') - U(M') |$$

We have only to show, that

$$\int_S | U(M') + U(M) | \, dS_M \quad \text{and} \quad \int_S | U_1(M') - U(M') | \, dS_M$$

are bounded. We have

$$\int_S | U_1(M') + U(M) | \, dS_M \leqq \int_{S_0} | d\nu(e_P) | \int_S \left| \frac{\cos\theta}{r^2} + \frac{\cos\theta'}{(r')^2} \right| dS_M .$$

The inner integral on the right is bounded by Lemma 2. Hence the iterated integral is bounded. Similarly it may be shown that

$$\int_S | U_1(M') - U(M') | \, dS_M$$

is bounded.

Let M, N on the surface S correspond to Q, P respectively on S_0.

LEMMA 3. *The integrals*

$$\int_{E_S} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2} dS_M \quad \textit{and} \quad \int_{E_S} \frac{\cos \sphericalangle (MN, n_M)}{\overline{MN}^2} dS_M$$

converge uniformly to the integrals

$$\int_{E_{S_0}} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2} dS_Q \quad \textit{and} \quad \int_{E_{S_0}} \frac{\cos \sphericalangle (QP, n_Q)}{\overline{QP}^2} dS_Q$$

respectively as τ approaches zero.

We may write

$$\left| \int_{E_{S_0}} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2} dS_Q - \int_{E_S} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2} dS_M \right| \leq J_1 + J_2$$

where

$$J_1 = \left| \int\limits_{E_{S_0} - E_{S_0} \cdot S_0(\delta, P)} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2} dS_Q - \int\limits_{E_S - E_S \cdot S(\delta, P)} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2} dS_M \right| ;$$

$$J_2 = \left| \int\limits_{E_{S_0} \cdot S_0(\delta, P)} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2} dS_Q - \int\limits_{E_S \cdot S(\delta, P)} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2} dS_M \right| .$$

Since these are surfaces of class Γ, the two integrals composing J_2 exist and approach zero uniformly with δ on account of the absolute continuity of the integral as a function of the set over which the integration is taken. Given $\epsilon > 0$ we may fix δ so that J_2 is less than $\epsilon/2$. Now let τ be less than δ so that N is in $S(\delta, P)$. The integrand of the second integral of J_1 converges uniformly to that of the first as τ approaches zero. Also dS_Q/dS_M approaches one uniformly as τ approaches zero. Hence for τ_1 sufficiently small we have $J_1 < \epsilon/2$ for all M, N, Q, P and all $\tau \leq \tau_1$, which proves the theorem. The other two integrals are treated similarly.

3. *On the solution of integral equations.* Let S_0 be a surface of the kind described in the preceding section and let $\{S\}$ be a normal family of surfaces inside or outside S_0 and including S_0. Let w be a regular closed curve [4] on S_0 and let σ be one part of S_0 enclosed by w. Denote by $q(P, w)$ the symmetric surface density of σ at the point P; i. e.

$$\begin{aligned} q(P, w) &= 1 \qquad \text{for } P \text{ inside } w, \\ &= 0 \qquad \text{for } P \text{ outside } w, \\ &= \psi/2\pi \ \text{ for } P \text{ on } w, \text{ where } \psi \text{ is the angle} \end{aligned}$$

between the forward and backward tangents to w at P. Define the function of regular curves with regular discontinuities

$$\nu(w) = \int_{S_0} q(P, w)\, d\nu(e_P) \tag{3.1}$$

Let $S(\tau)$ be a member of the family the least upper bound of whose normal distances from S_0 is τ, and let $w(\tau)$, $\sigma(\tau)$ be the sets of points on $S(\tau)$ corresponding to w, σ respectively on S_0. Write

$$v(M) = \int_{S_0} \frac{1}{MP}\, d\nu(e_P); \tag{3.2}$$

$$U(\tau, w) = \int_{\sigma(\tau)} U(M)\, dS_M; \tag{3.3}$$

$$V(\tau, w) = \int_{\sigma(\tau)} \frac{dv}{dn_M}\, dS_M = \int_{\sigma(\tau)} U_1(M)\, dS_M, \quad [\theta = \theta_M]. \tag{3.4}$$

Denoting approach to S_0 from T^+ and T^- by $\lim_{\tau \to 0+}$ and $\lim_{\tau \to 0-}$ respectively, we have [5]

$$\lim_{\tau \to 0^{\pm}} U(\tau, w) = \mp 2\pi\nu(w) + \int_{S_0} d\nu(e_P) \int_{\sigma} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2}\, dS_Q. \tag{3.5}$$

[4] (A), p. 498.

[5] (A), p. 502.

(3.6) $\lim_{\tau \to 0^\pm} V(\tau, w) = \mp 2\pi\nu(w) + \int_{S_0} d\nu(e_P) \int$

If $F(w)$ is a given completely additive functio with regular discontinuities, then in order to obtai determining $F(w)$ by means of the relation $F(\imath$ where $U(M)$ is given by (I) for M in T^+, we have

$$F(w) = -2\pi\nu(w) + \int_{S_0} d\nu(w_P) \int_\sigma \frac{\cos\cdot}{}$$

Putting

$$K(Q, P) = \frac{1}{2\pi} \frac{\cos \sphericalangle (QP, n_P)}{\overline{QP}^2} \qquad \Phi(w)$$

we obtain the equation

(3.7) $$\nu(w) = \Phi(w) + \int_{S_0} d\nu(w_P) \int_\sigma K$$

THEOREM 3.1. *The equation* (3.7) *has the* ε

(3.8) $$\nu(w) = \Phi(w) - \int_{S_0} d\Phi(w_P) \int_\sigma$$

where $k(Q, P)$ *is the resolvent kernel for the Fredh*

(3.9) $$h(Q) = f(Q) + \int_{S_0} K(Q, P) h($$

and where $f(Q)$ *is given by* (3.12) *below. The fun additive function of regular curves on* S_0 *with regu uniquely determined.*

The function $k(Q, P)$ is seen by its explicit ex below, to be continuous except when $Q = P$. More respect to Q and P over S_0, as follows incidentally below.

We show first that if $\nu(w)$ is a completely ad curves on S_0 with regular discontinuities and is a sol satisfies (3.8). Put

(3.10) $$\nu(w) = \Phi(w) + H(w).$$

It follows immediately from (3.7) that $H(w)$ Hence the derivative of $H(w)$ exists almost everyw

$$(3.11) \qquad H(w) = \int_{\sigma} h(R)\, dS_R,$$

where $h(R)$ is the derivative almost everywhere of $H(w)$. Substituting (3.10) in (3.7) we have

$$H(w) = \int_{S_0} d\Phi(w_P) \int_{\sigma} K(Q,P)\, dS_Q + \int_{S_0} dH(w_P) \int_{\sigma} K(Q,P)\, dS_Q.$$

Hence

$$h(Q) = \int_{S_0} K(Q,P)\, d\Phi(w_P) + \int_{S_0} K(Q,P) h(P)\, dS_P.$$

Putting

$$(3.12) \qquad f(Q) = \int_{S_0} K(Q,P)\, d\Phi(w_P)$$

we obtain the equation (3.9). But (3.9) has the unique [6] summable solution $h(Q)$ given by

$$(3.13) \qquad h(Q) = f(Q) - \int_{S_0} k(Q,R) f(R)\, dS_R.$$

where [7]

$$(3.14) \qquad k(Q,P) + K(Q,P) = \int_{S_0} k(Q,R) K(R,P)\, dS_R = \int_{S_0} K(Q,R) k(R,P)\, dS_R.$$

Substituting (3.13) in (3.11) we have

$$H(w) = \int_{\sigma} dS_Q \int_{S_0} K(Q,P)\, d\Phi(w_P) - \int_{\sigma} dS_Q \int_{S_0} k(Q,R)\, dS_R \int_{S_0} K(R,P)\, d\Phi(w_P).$$

By the first part of the lemma proved below we may change the order of integration to reduce this last equation to the form

$$H(w) = \int_{S_0} d\Phi(w_P) \int_{\sigma} dS_Q \{K(Q,P) - \int_{S_0} K(R,P) k(Q,R)\, dS_R\}.$$

Applying (3.14) we now have

$$H(w) = -\int_{S_0} d\Phi(w_P) \int_{\sigma} k(Q,P)\, dS_Q.$$

Substituting in (3.10), we obtain the equation (3.8). Similarly it may be

---

[6] (A), p. 507.

[7] See Appendix 1.

shown that $\nu(w)$ given by (3.8) is a completely additive function of curves and satisfies (3.7), which completes the proof of the theorem.

LEMMA 4. *The integrals*

$$\int_{S_0} |k(Q,P)|\, dS_Q \quad \text{and} \quad \int_{S_0} |k(Q,P)|\, dS_P$$

are bounded for all Q, P *on* S_0. *If* M, N *on* $S(\tau)$ *correspond respectively to* Q, P *on* S_0, *then*

$$\int_{S(\tau)} |k(M,N)|\, dS_M \quad \text{and} \quad \int_{S(\tau)} |k(M,N)|\, dS_N$$

are bounded uniformly as τ *approaches zero.*

We define the iterated kernels $K_i(Q,P)$ by the relation

$$K_{i+1}(Q,P) = \int_{S_0} K(Q,R) K_i(R,P)\, dS_R$$

where $K_0(R,P) = K(R,P)$. The kernel $K(Q,P)$ is not bounded. However $K_1(Q,P)$ is bounded and continuous except when $Q = P$ and $K_2(Q,P)$ is continuous [8] in Q and P. It may readily be shown that $K_1(M,N)$ is bounded uniformly as τ approaches zero.[9] Making use of this fact in conjunction with Lemma 3, we see that $K_i(M,N)$, $(i \geqq 2)$, changes continuously as τ approaches zero, the continuity being uniform in M, N. A proof of this last statement follows.

First, we prove the statement for $i = 2$. Let R on S_0 and G on $S(\tau)$ be corresponding points. Then for any M, N, Q, P

$$\left| \int_{S_0} K(Q,R) K_1(R,P)\, dS_R - \int_{S(\tau)} K(M,G) K_1(G,N)\, dS_G \right| \leqq J_1 + J_2$$

where

$$J_1 = \left| \int_{S_0 - S_0(\delta,Q) - S_0(\delta,P)} K(Q,R) K_1(R,P)\, dS_R - \int_{S(\tau) - S(\delta,Q) - S(\delta,P)} K(M,G) K_1(G,N)\, dS_G \right| ;$$

$$J_2 = \left| \int_{S_0(\delta,Q) + S_0(\delta,P)} K(Q,R) K_1(R,P)\, dS_R - \int_{S(\delta,Q) + S(\delta,P)} K(M,G) K_1(G,N)\, dS_G \right| .$$

Due to the uniform boundedness of $K_1(M,N)$ we may fix δ so that J_2 is as small as we like. Then we let τ be less than δ so that M and N are in $S(\delta, Q)$

---

[8] (A), p. 507. See also Kellogg, *Potential Theory*, p. 301.

[9] From the nature of the proof that $K_1(Q,P)$ is bounded, using the proof suggested in the footnote given on p. 507 of (A), the uniform boundedness of $K_1(M,N)$ is seen immediately. See Appendix 2.

and $S(\delta, P)$ respectively. Now the integrand of the second integral in J_1 converges uniformly to that of the first as τ approaches zero and dS_R/dS_G approaches one uniformly, which proves the statement. The proof for $i > 2$ is an obvious application of the method of mathematical induction.

Replacing $K(Q, P)$ by $\lambda K(Q, P)$ and substituting (3.9) into itself twice we obtain the equivalent [10] equation

$$h(Q) = f_2(Q) + \lambda^3 \int_{S_0} K_2(Q, P) h(P) \, dS_P \tag{3.15}$$

where

$$f_2(Q) = f(Q) + \lambda \int_{S_0} K(Q, P) f_1(P) \, dS_P$$

and

$$f_1(Q) = f(Q) + \lambda \int_{S_0} K(Q, P) f(P) \, dS_P.$$

The solution of (3.15) may be written

$$h(Q) = f_2(Q) - \lambda^3 \int_{S_0} k_2(Q, R; \lambda^3) f_2(R) \, dS_R \tag{3.16}$$

where $k_2(Q, R; \lambda^3)$ is continuous in Q and R, providing λ^3 is not a characteristic value for $k_2(Q, R; \lambda^3)$. Denoting by $k(Q, P; \lambda)$ the resolvent kernel for (3.9) where $K(Q, P)$ has been replaced by $\lambda K(Q, P)$ we have, from (3.13),

$$\begin{aligned} k(Q, R; \lambda) = {} & -K(Q, R) - \lambda K_1(Q, R) + \lambda^2 k_2(Q, R; \lambda^3) \\ & + \lambda^3 \int_{S_0} k_2(Q, P; \lambda^3) K(P, R) \, dS_P + \lambda^4 \int_{S_0} k_2(Q, P; \lambda^3) K_1(P, R) \, dS_P. \end{aligned} \tag{3.17}$$

Since $\lambda = 1$ is not a characteristic value for $K(Q, P)$ and the characteristic values of $K(Q, P)$ are real, $\lambda^3 = 1$ is not a characteristic value for $K_2(Q, P)$, and $k_2(Q, R; 1)$ is a continuous function of Q and R. From the known expressions for $k_2(Q, P; \lambda^3)$ in terms of the iterated kernels of index at least as great as two it is evident that $k_2(M, N; 1)$ converges uniformly to $k_2(Q, P; 1)$ as τ approaches zero. That $k(Q, P)$ and $k(M, N)$ have the properties stated in the lemma follows from (3.17) and the fact that $K(Q, P)$ and $K(M, N)$ have these properties.

From the solution (3.8) of the equation (3.7) we may obtain an explicit formula for the potential given by (I) in terms of $F(w)$. Substituting (3.8) in (I), we have

$$U(M) = \frac{1}{4\pi} \int_{S_0} g_n(M, P) \, dF(w_P) = \frac{1}{4\pi} \int_{S_0} g_n(M, P) \, dF(e_P) \tag{3.18}$$

[10] That (3.9) and (3.15) are equivalent if λ is not a characteristic value is shown in Goursat, *Cours d'Analyse Mathématique*, vol. 3, pp. 355-356.

where

$$(3.19)\quad g_n(M,P) = 2\int_{S_0} k(R,P)\frac{\cos \sphericalangle\,(MR, n_R)}{\overline{MR}^2}\,dS_R - 2\frac{\cos \sphericalangle\,(MP, n_P)}{\overline{MP}^2}.$$

The function of point sets $\nu(e)$ and the function of regular curves $\nu(w)$ are connected by the relations

$$(3.20)\quad \begin{cases} \nu(w) = \int_{S_0} q(P,w)\,d\nu(e_P) = \int_{S_0} q(P,w)\,d\nu(w_P) \\ \nu(e) = \int_{S_0} \phi(P,e)\,d\nu(w_P) = \int_{S_0} \phi(P,e)\,d\nu(e_P) \end{cases}$$

where $\phi(P,e)$ is the characteristic function of the set e. Similar relations exist[11] between $F(e)$ and $F(w)$, and also between the positive and negative variation functions $F_\pm(e)$, $F_\pm(w)$ and $\nu_\pm(e)$, $\nu_\pm(w)$. Since the function of point sets is completely additive, the corresponding function of regular curves is completely additive and conversely.

THEOREM 3.2. *If $U(M)$ is given by (I) for M in T^+, a necessary and sufficient condition that $\nu(e)$ be absolutely continuous is that $F(e)$ be absolutely continuous. Further, if in this situation $\nu'(P)$ is continuous, then $F'(P)$ is continuous and conversely, where $\nu'(P)$ and $F'(P)$ are the derivatives respectively of $\nu(e)$ and $F(e)$.*

From the equations corresponding to (3.20) for the positive and negative variation functions for $\nu(e)$ and $F(e)$ and $\nu(w)$ and $F(w)$ we have, denoting by $\bar\sigma$ the set of points composing w,

$$\begin{aligned} \nu_\pm(\sigma) &\leq \nu_\pm(w) \leq \nu_\pm(\sigma+\bar\sigma) \\ F_\pm(\sigma) &\leq F_\pm(w) \leq F_\pm(\sigma+\bar\sigma). \end{aligned}$$

Suppose $\nu(e)$ is absolutely continuous. Then $\nu_\pm(e)$ are absolutely continuous. Since $\nu_\pm(e)$ are also additive, we have

$$\nu_\pm(\sigma+\bar\sigma) = \nu_\pm(\sigma)$$

since $\bar\sigma$ is of zero superficial measures. Hence $\nu_\pm(w) = \nu_\pm(\sigma)$. From (3.7) we may write

$$F_\pm(\sigma) \leq F_\pm(w) \leq 2\pi\nu_+(\sigma) + 2\pi\nu_-(\sigma) + \int_{S_0} |\,d\nu(e_P)|\int_\sigma \frac{|\cos \sphericalangle\,(QP, n_P)|}{\overline{QP}^2}\,dS_Q$$

[11] See Bray and Evans, "A class of functions harmonic within the sphere," *American Journal of Mathematics*, vol. 49 (1927), pp. 158-159. The proofs given for functions of segments on the sphere apply equally well to bounded additive functions of regular curves on S_0.

where each term on the right is absolutely continuous. Therefore $F_{\pm}(e)$, and hence $F(e)$, are absolutely continuous. For the sufficiency of the condition stated in the theorem we use the same method of proof, but use equation (3.8) instead of (3.7).

We note that if $U(M)$ is given by (I) where $\nu(e) = \int_e \nu'(P)\,dS_P$ and if $\nu'(P)$ is continuous on S_0, then $U(M)$ takes on continuously its boundary values $F'(P)$. In fact

$$U(M) = \int_{S_0} \frac{\cos \sphericalangle (MP, n_P)}{\overline{MP}^2}\, \nu'(P)\, dS_P$$

and the statement follows as in the case of regular surfaces.[12]

THEOREM 3.3. *If $U(M)$ is given by (I), it may be written in T^+ in the form* (3.18) *where $g_n(M,P)$ is continuous in P and harmonic in M, for M in T^+, and is not negative; and*

$$F(w) = \lim_{\tau\to 0^+} U(\tau, w) = \lim_{\tau\to 0^+} \int_{\sigma(\tau)} U(M)\, dS_M$$

where M is on the surface $S(\tau)$ of the normal family inside S_0.

That $g_n(M,P)$ is harmonic in M and continuous in P for M in T^+ follows immediately[13] from the equation (3.19). We have only to show that $g_n(M,P)$ is not negative for M in T^+ and P on S_0. Suppose the contrary. Then for some point M_1 in T^+ and some point P_1 of S_0, $g_n(M,P)$ is negative. From the continuity of $g_n(M,P)$ in P there is a neighborhood ω_δ on S_0 containing P_1 such that $g_n(M_1,P) < 0$ for P in ω_δ. Suppose $F(w) = \int_\sigma F'(P)\,dS_P$ where $F'(P)$ is continuous on S_0, positive for P in ω_δ, and zero otherwise. Then

$$U(M_1) = \frac{1}{4\pi} \int_{\omega_\delta} g_n(M_1, P) F'(P)\, dS_P < 0.$$

But $U(M)$ cannot be negative anywhere in T^+ since it is the solution of the interior Dirichlet problem for continuous not-negative boundary values. This contradiction establishes the theorem.

THEOREM 3.4. *The function $U(M)$ given by (I) for M in T^+ is the difference of two not-negative functions harmonic in T^+.*

This theorem follows immediately from the equation (3.18) and the

[12] Kellogg, *Potential Theory*, pp. 167-168.
[13] See Appendix 3.

preceding theorem, together with the fact that the completely additive function $F(w)$ may be expressed as the difference of two not-negative functions.

THEOREM 3.5. *If $U(M)$ is given by (I) for M in T^+, a necessary and sufficient condition that $\nu(e)$ be absolutely continuous is that the absolute continuity of the integrals $\int_{\sigma(\tau)} U(M)\,dS_M$ be uniform as τ approaches zero.*

We prove first the necessity of the condition, supposing $\nu(e)$ to be absolutely continuous. $F(e)$ is absolutely continuous by the preceding theorem. It is sufficient to consider the case where $F(e)$, and therefore $U(M)$, is not negative. But

$$F(\sigma) = \lim_{\tau\to 0^+} \int_{\sigma(\tau)} U(M)\,dS_M.$$

Since $F(\sigma)$ is an absolutely continuous function of point sets, it has a derivative $F'(P)$ almost everywhere, and

$$(3.21) \qquad \begin{cases} F(\sigma) = \int_{\sigma} F'(Q)\,dS_Q \\ F'(Q) = \lim\limits_{\tau\to 0^+} U(M) \end{cases}$$

almost everywhere.[14] Suppose the absolute continuity of the integrals $\int_{\sigma(\tau)} U(M)\,dS_M$ is not uniform. Then there exists a denumerable sequence $\{S_i\}$ of surfaces of the family for which the absolute continuity of the integrals $\int_{\sigma_i} U(M)\,dS_M$ is not uniform, where σ_i denotes the region on S_i corresponding to σ on S_0. By the property (c) of normal families the sequence $\{S_i\}$ converges to some surface S^* of the family. We denote by $U_i(M^*)$ the function $U(M)\,dS_M/dS_{M^*}$ where M^* on S^* corresponds to M on S_i. Hence $\int_{\sigma_i} U(M)\,dS_M = \int_{\sigma^*} U_i(M^*)\,dS_{M^*}$, where σ^* on S^* corresponds to σ_i on S_i. By De la Vallée Poussin's converse to Vitali's Theorem, either $\{\lim\limits_{i\to\infty} U_i(M^*)\}$

---

[14] Due to the absolute continuity of F and ν, we may compute the derivatives of both members of (3.7), obtaining

$$\nu'(Q) = -(1/2\pi)F'(Q) + \int_{S_0} K(Q,P)\nu'(P)\,dS_P$$

where $\nu'(Q)$ and $F'(Q)$ are the derivatives almost everywhere of $\nu(w)$ and $F(w)$. Transposing, we obtain

$$F'(Q) = -2\pi\nu'(Q) + 2\pi\int_{S_0} K(Q,P)\nu'(P)\,dS_P = \lim_{\substack{\tau\to 0^+ \\ (M\to Q)}} U(M),$$

as is shown in (A), p. 511.

is not summable or $\lim_{i\to\infty}\int_{\sigma^*} U_i(M^*)\,dS_{M^*}$ is not equal to $\int_{\sigma^*}\{\lim_{i\to\infty} U_i(M^*)\}dS_{M^*}$. If S^* is S_0, neither of these conditions is possible as is shown by (3.21). But if S^* is not S_0, then $U(M)$ is continuous, and again neither condition can subsist.

To prove the sufficiency, we note that if the absolute continuity of the integrals $\int_{\sigma(\tau)} U(M)\,dS_M$ is uniform, then $F'(Q)$ is summable by Vitali's Theorem, and $\lim_{\tau\to 0^+}\int_{\sigma(\tau)} U(M)\,dS_M = \int_\sigma F'(Q)\,dS_Q$. Hence $F(e)$ is absolutely continuous and therefore $\nu(e)$ is absolutely continuous by Theorem 3.2.

4. *Necessary and sufficient conditions for representation as potentials.*

THEOREM 4.1. *A necessary and sufficient condition that $U(M)$ harmonic in T^+ be representable in the form (I) is that $\int_{S(\tau)} |U(M)|\,dS_M$ remain bounded over a normal family $\{S(\tau)\}$ within S_0 as τ approaches zero.*

The necessity of the condition is given by Theorem 2.3. To prove the sufficiency we note that by hypothesis the functions of point sets $F(\sigma(\tau))$, given by

$$F(\sigma(\tau)) = \int_{\sigma(\tau)} U(M)\,dS_M,$$

and therefore the corresponding functions of regular curves $F(w(\tau))$ defined on the surfaces $S(\tau)$, are of uniformly bounded total variation. Moreover $U(M)$ is uniformly continuous inside and on $S(\tau)$. From the equations corresponding to (3.8) formed for the surfaces of the family it follows that the functions $\nu(w(\tau))$ are of uniformly bounded total variation, and by Theorem 3.2 we may write $\nu(w(\tau)) = \int_{\sigma(\tau)} \nu'(M)\,dS_M$ where $\nu'(M)$ is continuous on $S(\tau)$. Consequently we may pick out a subsequence $\{S_k\}$ of surfaces of the family for which the functions $\nu_k(w')$ converge in the weak sense [15] to a completely additive function $\nu(w)$ defined on S_0, where the $\nu_k(w')$ represent the solutions of the equations corresponding to (3.7) for the surfaces S_k. Denoting by P_k a variable point on S_k, we have, for k sufficiently large,

$$U(M) = \int_{S_k} \frac{\cos \sphericalangle\,(MP_k, n_{P_k})}{\overline{MP_k}^2}\, d\nu_k(w'_{P_k}).$$

[15] J. Radon, "Über die Randwertaufgaben beim Logarithmischen Potential," *Wiener Akademie Sitzungsberichte*, vol. 128 (1919), IIa, p. 1153. The methods employed by Radon for a certain class of curves may be extended readily to apply to normal families of surfaces. See Appendix 4.

In fact, the $\nu_k(e)$ is the integral of a continuous $\nu'(P_k)$ and the right-hand member is harmonic within S_k and takes on continuously the boundary values $U(M_k)$ as M tends to M_k on S_k. But there can be only one such function. Consequently that function is $U(M)$. From the weak convergence of the functions $\nu_k(w')$ to $\nu(w)$ defined on S_0, we now have

$$U(M)=\int_{S_0}\frac{\cos\sphericalangle(MP,n_P)}{\overline{MP}^2}\,d\nu(w_P)=\int_{S_0}\frac{\cos\sphericalangle(MP,n_P)}{\overline{MP}^2}\,d\nu(e_P)$$

which is what was to be shown.

LEMMA 5. *The function $g_n(M,P)$ given by (3.18) is positive for M in T^+ and P on S_0.*

By Theorem 3.3, $g_n(M,P)$ is not-negative. Suppose there exists a point M_1 of T^+ and a point P_1 of S_0 such that $g_n(M_1,P_1)$ is zero. Then $g_n(M,P_1)$ is identically zero, since $g_n(M,P)$ is harmonic in M and not-negative. For any given completely additive function of regular curves with regular discontinuities, $F(w)$, we may solve the equation (3.7) for $\nu(w)$ so that $U(M)$ is given by (3.18) and

$$F(w)=\lim_{\tau\to 0^+}\int_{\sigma(\tau)}U(M)\,dS_M.$$

Define $F(e)$ so that

$$F(e)=1 \text{ if } e \text{ contains } P_1.$$
$$F(e)=0 \text{ otherwise.}$$

From (3.17) we now have $U(M)\equiv 0$. Hence

$$F(w)=\lim_{\tau\to 0^+}\int_{\sigma(\tau)}U(M)\,dS_M=0$$

for all w, and therefore $F(e)=0$ for all e, which contradicts the assumption that $F(e)=1$ for all e containing P_1.

THEOREM 4.2. *A necessary and sufficient condition*[16] *that $U(M)$ harmonic in T^+ be given by (I) is that $U(M)$ be the difference of two not-negative functions harmonic in T^+.*

The necessity of the condition is given by Theorem 3.4. For the suffi-

[16] De la Vallée Poussin (*loc. cit.*, p. 199) proves that a necessary and sufficient condition that the harmonic function U defined in T^+ be the difference of two positive harmonic functions is that $\int_S |U|\,dS$ be bounded over a normal family of surfaces in T^+ and arbitrarily near S_0. This theorem follows directly from Theorems 4.1 and 4.2.

ciency we need only prove the theorem for $U(M)$ not-negative and harmonic in T^+. If $U(M)$ vanishes anywhere in T^+, then it vanishes identically, and the theorem is trivial. Hence, we shall suppose that $U(M)$ is positive. Consider a fixed point $\bar{M}$ of T^+ and a normal family $S(\tau)$ of surfaces inside S_0, where $\tau_1 \geqq \tau \geqq 0$ and τ_1 is small enough so that $\bar{M}$ is inside all the surfaces. We form the functions $g_n(\bar{M}, N)$ corresponding to $g_n(\bar{M}, P)$ where N is a point of $S(\tau)$. From the form of $g_n(\bar{M}, P)$ given by (3. 19) and from Lemma 4 it is easily seen that $g_n(\bar{M}, N)$ converges uniformly [17] to $g_n(\bar{M}, P)$ as N on $S(\tau)$ approaches P on S_0 along n_P and that the convergence is uniform in τ and P. There exists a number $\epsilon > 0$ such that

$$g_n(\bar{M}, N) \geqq \epsilon$$

for all N of $S(\tau)$ and all τ; for, supposing the contrary, we can find a sequence of points N_i on surfaces of the family, having a limit point $\bar{N}$ on one of the surfaces such that $g_n(\bar{M}, \bar{N})$ equals zero. This is impossible by Lemma 5. Defining the functions $F(w(\tau))$ on $S(\tau)$ by the relation

$$F(w(\tau)) = \int_{S(\tau)} U(M)\, dS_M$$

we have for every τ in $\tau_1 \geqq \tau \geqq 0$

$$U(\bar{M}) = \frac{1}{4\pi} \int_{S(\tau)} g_n(\bar{M}, N)\, dF(w(\tau)_N).$$

In fact $F(\sigma(\tau))$ is the integral of a continuous function $F'(N)$, and $F'(N) = \lim\limits_{M \to N} U(M)$. That $U(\bar{M})$ may be written in this form follows from Theorem 3. 3 and the fact that inside $S(\tau)$, $U(\bar{M})$ is given by

$$U(M) = \int_{S(\tau)} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2}\, dv(w(\tau)_N)$$

as was noted in the proof of Theorem 4. 1. Now $F(w(\tau))$ is positive since $U(M)$ is positive. Hence

$$U(\bar{M}) \geqq \frac{\epsilon}{4\pi} F(S(\tau)).$$

Therefore the functions $F(S(\tau))$ remain bounded uniformly over the normal family inside S_0. The theorem now follows immediately from Theorem 4. 1.

THEOREM 4. 3. *A necessary and sufficient condition that $v(M)$ harmonic in T^-, except at infinity, be representable in T^- by* (3. 2) *is that*

---

[17] Appendix 5.

(a) *$v(M)$ approach zero continuously at infinity and*

(b) $\int_{S(\tau)} |U_1(M)|\, dS_M = \int_{S(\tau)} \left| \frac{dv}{dn_M} \right| dS_M$ *remain bounded over a normal family $S(\tau)$ outside S_0 as τ approaches zero.*

The necessity of (a) is immediate. The necessity of (b) is given by Theorem 2. 3. Before proving the sufficiency we consider (3. 6) in the form

$$G(w) = 2\pi\nu(w) + \int_{S_0} d\nu(e_P) \int_\sigma \frac{\cos \sphericalangle\, (QP, n_Q)}{\overline{QP}^2}\, dS_Q;$$

or

$$\nu(w) = \Psi(w) + \int_{S_0} d\nu(e_P) \int_\sigma K(P, Q)\, dS_Q \tag{4. 1}$$

where

$$\Psi(w) = \frac{1}{2\pi}\, G(w) = \lim_{\tau \to 0-} \frac{1}{2\pi} \int_{\sigma(\tau)} \frac{dv}{dn_M}\, dS_M.$$

We may solve this equation in a manner similar to the one followed in solving (3. 7). We may write the solution in the form

$$\nu(w) = \Psi(w) - \int_{S_0} d\Psi(w_P) \int_\sigma k(P, R)\, dS_R. \tag{4. 2}$$

Given $v(M)$, in order to prove the sufficiency of the conditions (a) and (b), we form the functions $\Psi(w(\tau)) = \int_{\sigma(\tau)} U_1(M)\, dS_M$. By hypothesis, these functions are of uniformly bounded total variation. From the equations corresponding to (4. 2) formed for the surfaces $S(\tau)$ it is seen that the functions $\nu(w(\tau))$ have the same property. Hence, as in the proof of Theorem 4. 1, we may pick out subsequences $\nu_k(w')$ defined on surfaces S_k of the family outside S_0 converging in the weak sense to a completely additive function $\nu(w)$ defined on S_0. From (a) it follows that $v(M)$ is regular [18] at infinity. Since $U_1(M)$ is continuous outside S_k and takes on continuous boundary values on S_k, it follows from (4. 2) that $\nu_k(e)$ is the integral of a continuous function. Hence for k sufficiently large, we may write as for regular surfaces [19]

---

[18] It follows immediately from a theorem given by Poincaré in his *Théorie du Potentiel Newtonien*, p. 210, that we may write

$$v(M) = v(x, y, z) = K + H_0/R + H_1/R^3 + \cdots + H_n/R^{2n+1} + \cdots$$

valid outside a sphere of sufficiently large radius; where K is a constant; H_i is a spherical harmonic of degree i in x, y, z; and $R = \sqrt{x^2 + y^2 + z^2}$. We note at once that $K = 0$, and hence $v(M)$ approaches zero canonically at infinity.

[19] Kellogg, *Potential Theory*, p. 311.

$$v(M) = \int_{S_k} \frac{1}{MP_k} d\nu_k(w'_{P_k})$$

for the right-hand member is harmonic outside S_k and is regular at infinity and its normal derivative takes on continuously[20] the boundary values $U_1(M_k)$ as M tends to M_k on S_k. But there can be only one such function.[21] Consequently that function is $v(M)$. From the weak convergence, we now have

$$v(M) = \int_{S_0} \frac{1}{MP} d\nu(w_P) = \int_{S_0} \frac{1}{MP} d\nu(e_P),$$

which was to be shown.

THEOREM 4. 4. *A necessary and sufficient condition that $v(M)$ harmonic in T^+ be representable in T^+ in the form* (3. 2) *is that $\int_{S(\tau)} | U_1(M) | dS_M$ remain bounded over a normal family $S(\tau)$ inside S_0 as τ approaches zero.*

The necessity of the condition is given by Theorem 2. 3. Incidentally, for $v(M)$ harmonic in T^+ we have $G(S(\tau)) = \int_{S(\tau)} \frac{dv}{dn_M} dS_M = 0$. Before proving the sufficiency we first consider the equation corresponding to (3. 6). This equation is

$$\nu(w) = \Psi(w) - \int_{S_0} d\nu(w_P) \int_\sigma K(P, Q) dS_Q \tag{4. 3}$$

where

$$\Psi(w) = -\frac{1}{2\pi} \lim_{\tau \to 0+} \int_{\sigma(\tau)} \frac{dv}{dn_M} dS_M.$$

The value $\lambda = -1$ is a characteristic value for the kernels $\lambda K(Q, P)$ and $\lambda K(P, Q)$ and the equation (4. 3) can be satisfied if and only if $\Psi(S_0)$ is zero. The mass function $\nu(w)$ is then determined except for an arbitrary additive term of the form $C \int_\sigma \phi_2(P) dS_P$, where C is a constant and $\phi_2(P)$ is a solution of the homogeneous equation[22]

$$\phi(P) = - \int_{S_0} \phi(Q) K(Q, P) dS_Q.$$

From the resolvent kernel $k(P, Q)$ we can find a particular solution of (4. 3).

[20] *Ibid.*, p. 164.
[21] *Ibid.*, p. 213.
[22] (A), p. 505. (An error in printing occurs here. The λ should be replaced by $-\lambda$).

We first reduce (4.3) to the Fredholm form. We put $\nu(w) = \Psi(w) - H(w)$. From (4.3) it is seen that $H(w)$ is absolutely continuous, and hence we may write

$$H(w) = \int_\sigma h(R)\, dS_R$$

where $h(R)$ is the derivative almost everywhere of $H(w)$. Proceeding as in the proof of Theorem 4.1 we obtain the Fredholm equation

$$h(Q) = f(Q) - \int_{S_0} K(P, Q) h(P)\, dS_P \tag{4.4}$$

where

$$f(Q) = \int_{S_0} K(P, Q)\, d\Psi(w_P).$$

We write

$$k(P, Q; \lambda) = \frac{A(P, Q)}{\lambda + 1} + B(P, Q; \lambda)$$

where $A(P, Q)$ is the residue at the pole $\lambda = -1$ and is continuous.[23] A particular solution of (4.4) is now given by

$$h(Q) = f(Q) + \int_{S_0} B(P, Q; -1) f(P)\, dS_P$$

or

$$h(Q) = f(Q) + \int_{S_0} B(P, Q; -1) f(P)\, dS_P - \int_{S_0} A(P, Q) f(P)\, dS_P \tag{4.5}$$

since

$$\int_{S_0} A(P, Q) f(P)\, dS_P = 0.^{24}$$

From (4.5) we have

$$\begin{aligned} H(w) = &\int_\sigma dS_Q \int_{S_0} K(P, Q)\, d\Psi(w_P) \\ &+ \int_\sigma dS_Q \int_{S_0} B(R, Q; -1)\, dS_R \int_{S_0} K(P, R)\, d\Psi(w_P) \\ &- \int_\sigma dS_Q \int_{S_0} A(R, Q)\, dS_R \int_{S_0} K(P, R)\, d\Psi(w_P). \end{aligned}$$

The first part of the lemma proved below permits us to change the order of integration in this equation so that we may write

$$\begin{aligned} H(w) = \int_{S_0} d\Psi(w_P) \int_\sigma dS_Q \{K(P, Q) + &\int_{S_0} K(P, R) B(R, Q; -1)\, dS_R \\ &- \int_{S_0} K(P, R) A(R, Q)\, dS_R\}. \end{aligned}$$

[23] Kellogg, *Potential Theory*, p. 308.
[24] Kellogg, *loc. cit.*, p. 298.

From (3.14) we obtain the relation

$$K(P,Q)+B(P,Q;\lambda) = \int_{S_0} K(P,R)A(R,Q)\,dS_R - \int_{S_0} K(P,R)B(R,Q;\lambda)\,dS_R.$$

Substituting in the previous equation, we now have

$$H(w) = -\int_{S_0} d\Psi(w_P)\int_\sigma B(P,Q;-1)\,dS_Q$$

or

$$\text{(4.6)} \qquad \nu(w) = \Psi(w) + \int_{S_0} d\Psi(w_P)\int_\sigma B(P,Q;-1)\,dS_Q$$

as a particular solution of (4.3)

LEMMA 6. *The integrals*

$$\int_{S_0} |B(P,Q;-1)|\,dS_P \quad \textit{and} \quad \int_{S_0} |B(P,Q;-1)|\,dS_Q$$

are bounded for all Q, P. *The integrals*

$$\int_{S(\tau)} |B(N,M;-1)|\,dS_N \quad \textit{and} \quad \int_{S(\tau)} |B(N,M;-1)|\,dS_M$$

are bounded uniformly as τ *approaches zero.*

We may write

$$k_2(Q,P;\lambda^3) = \frac{A_2(Q,P)}{\lambda^3+1} + B_2(Q,P;\lambda^3)$$

where $A_2(Q,P)$ is continuous in Q, P and $B_2(Q,P;\lambda^3)$ is a power series in (λ^3+1) uniformly convergent in the neighborhood of $\lambda=-1$ and with coefficients which are continuous[25] in Q and P. Moreover, $B_2(M,N;-1)$ and $A_2(M,N)$ are bounded as τ approaches zero since $k_2(M,N;\lambda^3)$ changes continuously with τ. Substituting in (3.17) and equating like powers of $(\lambda+1)$, we obtain for $\lambda=-1$

$$\text{(4.7)} \quad B(Q,R;-1) = -K(Q,R)+K_1(Q,R)-\tfrac{1}{3}A_2(Q,R)+B_2(Q,R;-1) + \tfrac{2}{3}\int_{S_0} A_2(Q,P)K(P,R)\,dS_P - \int_{S_0} A_2(Q,P)K_1(P,R)\,dS_P - \int_{S_0} B_2(Q,P;-1)K(P,R)\,dS_P + \int_{S_0} B_2(Q,P;-1)K_1(P,R)\,dS_P.$$

---

[25] Kellogg, *loc. cit.*, p. 294.

Since

$$A_2(Q,R) = -\int_{S_0} K_2(Q,P)A_2(P,R)\,dS_P$$
$$= -\int_{S_0} K(Q,P)A_2(P,R)\,dS_P = +\int_{S_0} K_1(Q,P)A_2(P,R)\,dS_P$$

we reduce (4. 7) to the form

(4. 8) $B(Q,R;-1)$
$$= -K(Q,R) + K_1(Q,R) + B_2(Q,R;-1) - 2A_2(Q,R)$$
$$-\int_{S_0} B_2(Q,P;-1)K(P,R)\,dS_P + \int_{S_0} B_2(Q,P;-1)K_1(P,R)\,dS_P.$$

The lemma is proved by treating (4. 8) in precisely the same manner as we treated (3. 17) in proving Lemma 4.

We continue now with the proof of Theorem 4. 4. By hypothesis the functions $\Psi(w(\tau)) = -\dfrac{1}{2\pi}\displaystyle\int_{\sigma(\tau)} \frac{dv}{dn_M}\,dS_M$ are of uniformly bounded total variation. Since $\Psi(S(\tau)) = 0$, the equations corresponding to (4. 3) formed for the surfaces $S(\tau)$ may be solved for the mass functions $\bar{\nu}(w(\tau))$ defining on $S(\tau)$ the functions[26] $\Psi(w(\tau))$. The functions $\bar{\nu}(w(\tau))$, particular solutions of the equations corresponding to (4. 3), given by the equations of the form (4. 6) are of uniformly bounded total variation. Hence we may select a subsequence $\bar{\nu}_k(w')$ defined on the surfaces S_k and converging in the weak sense to $\bar{\nu}(w)$ defined on S_0. These functions $\bar{\nu}_k(w')$ satisfy the equations of the form (4. 3), but so also do the functions

$$\nu_k(w') = \bar{\nu}_k(w') + C_k\int_{\sigma'} \phi_2(P_k)\,dS_{P_k}$$
$$= \bar{\nu}_k(w') + \nu^*_k(w')$$

where w', σ', P_k on S_k correspond respectively to w, σ, P on S_0; C_k is an arbitrary constant; and $\phi_2(P_k)$ is a solution of the homogeneous equation $\phi(P_k) = -\int_{S_k} \phi(M)K(M,P_k)\,dS_M$. The potential of the form (3. 2) due to the mass function $\nu^*_k(w')$ reduces inside S_k to a constant[27] K_k, where

$$K_k = \int_{S_k} \frac{1}{MP_k}\,d\nu^*_k(w'_{P_k}) = C_k\int_{S_k} \frac{1}{MP_k}\,\phi_2(P_k)\,dS_{P_k}.$$

By properly choosing the constants K_k, we have

[26] (A), p. 505.
[27] (A), p. 509 (footnote).

$$v(M) - K_k = \int_{S_k} \frac{1}{MP_k}\, d\nu_k(w'_{P_k})$$

for M inside S_k; for, as in the previous theorem, $\nu_k(w')$ is the integral of a continuous function and the normal derivative of the right-hand member takes on the same boundary values as does the normal derivative of $v(M)$ and therefore the right-hand member differs from $v(M)$ by a constant.[28] Since $v(M)$ is given and the functions $\bar{\nu}_k(w')$ converge in the weak sense to $\bar{\nu}(w)$ defined on S_0, we have for M in T^+

$$v(M) = \int_{S_0} \frac{1}{MP}\, d\bar{\nu}(w_P) + K$$

where $K = \lim_{k\to\infty} K_k$. But inside T^+ we may write

$$K = B \int_{S_0} \frac{1}{MP}\, \phi_2(P)\, dS_P,$$

where B is a constant, due to the fact that

$$v_0(M) = \int_{S_0} \frac{1}{MP}\, \phi_2(P)\, dS_P \neq 0$$

(for otherwise, $v_0(M)$ being a conductor potential, we have $\phi_2(P) \equiv 0$, which is impossible since $\phi_2(P)$ is a non-zero solution of the homogeneous equation). Defining $\nu(w)$ by the relation

$$\nu(w) = \bar{\nu}(w) + B \int_{\sigma} \phi_2(P)\, dS_P$$

we now have

$$v(M) = \int_{S_0} \frac{1}{MP}\, d\nu(w_P) = \int_{S_0} \frac{1}{MP}\, d\nu(e_P)$$

which establishes the theorem.

THEOREM 4.5. *A necessary and sufficient condition that $U(M)$ harmonic in T^-, except at infinity, be representable in T^- in the form (I) is that*

(a) *$U(M)$ approach zero continuously at infinity and* $\int_{\Omega} \frac{dU}{dn}\, dS = 0$, *where Ω is a sphere of radius $r > r_0$ sufficiently large, and*

(b) *that* $\int_{S(\tau)} |U(M)|\, dS_M$ *remain bounded uniformly over a normal family $S(\tau)$ outside S_0 as τ approaches zero.*

[28] Kellogg, *Potential Theory*, p. 213.

The necessity of (a) is well known. The necessity of (b) is given by Theorem 2. 3. To prove the sufficiency we must be able to solve the integral equations corresponding to (3. 5) for the surfaces $S(\tau)$. These equations take the form

$$\nu(w(\tau)) = \Phi(w(\tau)) - \int_{S_0} d\nu(e_N) \int_{\sigma(\tau)} K(M,N)\, dS_M \tag{4. 9}$$

where

$$\Phi(w(\tau)) = \frac{1}{2\pi} \int_{\sigma(\tau)} U(M)\, dS_M.$$

A sufficient condition that these equations have solutions is that

$$\int_{S(\tau)} \phi_2(M)\, d\Phi(e_M) = \frac{1}{2\pi} \int_{S(\tau)} \phi_2(M)\, U(M)\, dS_M = 0$$

where $\phi_2(M)$ is a solution of the homogeneous equation

$$\phi(N) = -\int_{S(\tau)} \phi(M) K(M,N)\, dS_M.$$

But the potential of the form (3. 2) for $\nu(w(\tau)) = \int_{S(\tau)} \phi_2(M)\, dS_M$ may be expressed as a conductor potential, and hence we may take $\phi_2(M)$ as the density of the distribution producing the conductor potential, which we denote by $V(M)$. Now let Γ be a sphere of sufficiently large radius, a, and fixed center. We have

$$\int_{S(\tau)+\Gamma} U \frac{dV}{dn^-}\, dS - \int_{S(\tau)+\Gamma} V \frac{dU}{dn^-}\, dS = 0.$$

Since $U(M)$ vanishes at infinity like $1/r^2$ and dU/dn^- vanishes like $1/r^3$,[29] we may let a become infinite, obtaining

---

[29] As we noted in the footnote to Theorem 4. 3, we may write

$$U(M) = U(x, y, x) = H_0/R + H_1/R^3 + \cdots + H_n/R^{2n+1} + \cdots$$

uniformly convergent outside a sphere of radius sufficiently large, where

$$R = \sqrt{x^2 + y^2 + z^2}.$$

Let Ω be a sphere of radius r large enough so that the above series converges uniformly outside and on Ω. On Ω we have

$$U(M) = H_0/r + H_1/r^3 + \cdots + H_n/r^{2n+1} + \cdots$$

and

$$\frac{dU}{dn} = -\frac{\partial U}{\partial r} = -\frac{\partial U}{\partial x}\frac{x}{r} - \frac{\partial U}{\partial y}\frac{y}{r} - \frac{\partial U}{\partial z}\frac{z}{r}.$$

But

$$\frac{\partial}{\partial r}\left(\frac{H_n}{r^{2n+1}}\right) = \frac{r^{2n+1}(H_n''/r) - (2n+1)H_n r^{2n}}{r^{4n+2}} = \frac{-H'_n(x, y, z)}{r^{2n+2}}.$$

$$\int_{S(\tau)} U \frac{dV}{dn^-} dS - \int_{S(\tau)} V \frac{dU}{dn^-} dS = 0.$$

But on $S(\tau)$, dV/dn^- vanishes identically. Moreover

$$dV/dn_M^- - dV/dn_M^+ = 4\pi\phi_2(M).$$

Hence we have, since $V = 1$ on $S(\tau)$ and $\int_{S(\tau)} \frac{dU}{dn_M^-} dS_M = 0$,

$$\int_{S(\tau)} U \frac{dV}{dn_M^-} dS_M = 4\pi \int_{S(\tau)} \phi_2(M) U(M) dS_M = 0.$$

Therefore the equations (4. 9) may be solved. Consequently we may write $U(M)$ outside $S(\tau)$ in the form

$$U(M) = \int_{S(\tau)} \frac{\cos \sphericalangle (MN, n_N)}{\overline{MN}^2} d\nu(w(\tau)_N)$$

where $\nu(w(\tau))$ is the solution of (4. 9); for the right-hand member is harmonic outside $S(\tau)$, vanishes at infinity, and takes on the same values on $S(\tau)$ as $U(M)$, and is therefore identically equal to $U(M)$. The solutions of (4. 9) contain arbitrary additive constants, but we may take these constants to be zero, since they add nothing to the potentials given in the form (I) for M in T^-. To complete the proof we apply the weak convergence analysis as in the preceding theorems.

APPENDIX.

1.

We derive the equation (3. 14) by making use of the formula for $k(Q, P)$ in terms of K, K_1, and k_2; namely, the equation (3. 17). For $\lambda = 1$ the equation (3. 17) becomes

---

where H'' and H' are spherical harmonics of order n. Hence

$$\partial U/\partial n = H_0/r^2 + H'_1/r^4 + \cdots + H'_n/r^{2n+2} + \cdots.$$

Therefore $\lim_{r\to\infty} \int_\Omega \frac{dU}{dn} dS = 4\pi H_0$. But from our assumption concerning the behavior of $U(M)$ at infinity, it follows that $\lim_{r\to\infty} \int_\Omega \frac{dU}{dn} dS = 0$. Hence $H_0 = 0$, and therefore $U(M)$ and dU/dn vanish at infinity like $1/r^2$ and $1/r^3$ respectively.

$$(1.1') \qquad k(Q,R) = -K(Q,R) - K_1(Q,R) + k_2(Q,R) + \int_{S_0} k_2(Q,x)\{K(x,R) + K_1(x,R)\}dS_x.$$

Since $K_2(Q,P)$ is continuous in Q and P, $K_2(Q,P)$ satisfies the Volterra relations

$$(1.2') \qquad \int_{S_0} k_2(Q,R)K_2(R,P)\,dS_R = \int_{S_0} k_2(R,P)K_2(Q,R)\,dS_R = k_2(Q,P) + K_2(Q,P).$$

From (1.1′) we have

$$\begin{aligned}\int_{S_0} k(Q,R)K(R,P)\,dS_R &= -\int_{S_0} K(Q,R)K(R,P)\,dS_R \\ &\quad - \int_{S_0} K_1(Q,R)K(R,P)\,dS_R + \int_{S_0} k_2(Q,R)K(R,P)\,dS_R \\ &\quad + \int_{S_0} k_2(Q,x)\,dS_x \int_{S_0} \{K(x,R)K(R,P) + K_1(x,R)K(R,P)\}dS_R \\ &= -K_1(Q,P) - K_2(Q,P) + \int_{S_0} k_2(Q,x)\,dS_x\{K(x,P) + K_1(x,P)\} \\ &\quad + \int_{S_0} k_2(Q,x)K_2(x,P)\,dS_x.\end{aligned}$$

Replacing the last integral by its value given by (1.2′), we have

$$\begin{aligned}(1.3') \qquad \int_{S_0} k(Q,R)K(R,P)\,dS_R &= -K(Q,P) + k_2(Q,P) \\ &\quad + \int_{S_0} k_2(Q,x)\{K(x,P) + K_1(x,P)\}dS_x \\ &= k(Q,P) + K(Q,P)\end{aligned}$$

by (1.1′).

Since

$$-k_2(Q,R) = K_2(Q,R) + K_5(Q,R) + K_8(Q,R) + \cdots,$$

this series being uniformly convergent for Q, R on S_0; and since

$$\int_{S_0} K_i(Q,x)K_j(x,R)\,dS_x = \int_{S_0} K_j(Q,x)K_i(x,R)\,dS_x = K_{i+j+1}(Q,R)$$

it follows that

$$(1.4') \qquad \int_{S_0} k_2(Q,x)K_i(x,R)\,dS_x = \int_{S_0} K_i(Q,x)k_2(x,R)\,dS_x \quad (i = 0, 1, \cdots).$$

By (1. 1′)

$$k(R,P)=-K(R,P)-K_1(R,P)+k_2(R,P)$$
$$+\int_{S_0}k_2(R,x)\{K(x,P)+K_1(x,P)\}dS_x.$$

Making use of (1. 4′) we write this equation in the equivalent form

$$(1.\,5')\qquad k(R,P)=-K(R,P)-K_1(R,P)+k_2(R,P)$$
$$+\int_{S_0}k_2(x,P)\{K(R,x)+K_1(R,x)\}dS_x.$$

Hence we have

$$\int_{S_0}k(R,P)K(Q,R)\,dS_R$$
$$=-K_1(Q,P)-K_2(Q,P)+\int_{S_0}k_2(x,P)K(Q,x)\,dS_x$$
$$+\int_{S_0}K(Q,R)\,dS_R\int_{S_0}k_2(x,P)\{K(R,x)+K_1(R,x)\}dS_x$$
$$=-K_1(Q,P)-K_2(Q,P)+\int_{S_0}k_2(x,P)\{K(Q,x)+K_1(Q,x)\}dS_x$$
$$+\int_{S_0}k_2(x,P)K_2(Q,x)\,dS_x.$$

By (1. 2′) this last equation may be written

$$(1.\,6')\qquad \int_{S_0}k(R,P)K(Q,R)\,dS_R$$
$$=-K_1(Q,P)+k_2(Q,P)+\int_{S_0}k_2(x,P)\{K(Q,x)+K_1(Q,x)\}dS_x$$
$$=k(Q,P)+K(Q,P)$$

from (1. 5′). Combining (1. 3′) and (1. 6′) we obtain the Volterra relations (3. 14), which was to be shown.

2.

LEMMA 1a. *$K_1(Q,P)$ is bounded.*

We suppose δ to be a positive constant, and consider two cases.

Case 1. $\overline{QP}\geqq 2\delta>0$. We may write

$$4\pi^2K_1(Q,P)=4\pi^2\int_{S_0-S_0(\delta,Q)-S_0(\delta,P)}K(Q,R)K(R,P)\,dS_R+4\pi^2\int_{S_0(\delta,Q)}+4\pi^2\int_{S_0(\delta,P)}.$$

We denote the three terms on the right by I_1, I_2, and have

$$|I_1| \leq \frac{1}{\delta^4} m(S_0)$$

$$|I_2| \leq \frac{1}{\delta^2} \int_{S_0(\delta,Q)} \frac{|\cos \sphericalangle (QR, n_R)|}{\overline{QR}^2}.$$

$$|I_3| \leq \frac{1}{\delta^2} \int_{S_0(\delta,P)} \frac{|\cos \sphericalangle (RP, n_P)|}{\overline{RP}^2}$$

It is obvious that $|I_1|$ is bounded. $|I_2|$ and $|I_3|$ are surfaces of class Γ.

Case 2. $\overline{QP} < 2\delta$. We let $\delta' = 4\delta$ and write

$$4\pi^2 K_1(Q, P) = 4\pi^2 \int_{S_0 - S_0(\delta', P)} K(Q, R) K(R, P)$$

We denote the terms on the right by J_1 and J_2 respe

$$|J_1| \leq \frac{1}{64\delta^4} m(S_0)$$

$$|J_2| \leq \int_{S_0(\delta', P)} \left| \frac{\cos \sphericalangle (QR, n_R)}{\overline{QR}^2} \cdot \frac{\cos \sphericalangle (RI}{\overline{RP}^2}\right.$$

Since $|J_1|$ is obviously bounded, we have only to prc

Let arc PR measured along the curve of interse determined by n_P and R be s. Let n_P be the z-axis S_0 at P be the xy-plane. Let R have cylindrical co similar manner set up a system of cylindrical coördina be the coördinates of R in this system. Denote by in the tangent plane at P. Let $\bar{s}$ be the arc QR mea intersection of S_0 and the plane determined by n_Q enough so that

$$(2.1') \quad \begin{cases} |\sphericalangle (QR, n_R)| \leq \pi; \quad |\sphericalangle (QR, n_Q)| \leq \\ \rho'/\rho'_1 < 2 \\ s < 2\rho \\ ds < 2d\rho \\ \bar{s} < 2\rho'_1 \\ d\bar{s} < 2d\rho'_1. \end{cases}$$

Now we have

$$
\begin{aligned}
&|\cos \sphericalangle (RP, n_P)| = z/RP < 4\gamma\rho \text{ (see the 9th of the inequalities (2.1))}\\
&|\cos \sphericalangle (QR, n_Q)| = z'/RQ < 4\gamma\rho'_1\\
&|\sphericalangle (QR, n_R) - \sphericalangle (QR, n_Q)| \leq (n_Q, n_R)|\\
&|\cos \sphericalangle (QR, n_R) - \cos \sphericalangle (QR, n_Q)| \leq |\sphericalangle (n_Q, n_R)| < \gamma\overline{QR}.
\end{aligned}
$$

Hence

$$
\begin{aligned}
|\cos \sphericalangle (QR, n_R)| &< |\cos \sphericalangle (QR, n_Q)| + \gamma\overline{QR}\\
&\leq 4\gamma\rho'_1 + \gamma\overline{QR}.
\end{aligned}
$$

Therefore

$$
\left|\frac{\cos \sphericalangle (RP, n_P)}{\overline{RP}^2}\right| < 4\gamma\rho/RP^2 \leq 4\gamma/\rho
$$

$$
\left|\frac{\cos \sphericalangle (QR, n_R)}{\overline{QR}^2}\right| < 4\gamma\rho'_1/\overline{QR}^2 + \gamma/\overline{QR} \leq 5\gamma/\rho'_1 \leq 10\gamma/\rho'.
$$

We have now

$$
(2.2') \qquad |J_2| < 40\gamma^2 \int_{S_0(\delta',P)} \frac{dS_R}{\rho\rho'}.
$$

Let $C(\delta', P)$ be the circle with center P and radius δ' and lying in the tangent plane to S_0 at P. We have $|dS_R| < 2dS'$ where dS' is the projection of dS_R on the tangent plane at P. We have finally

$$
|J_2| < 80\gamma^2 \int_{C(\delta',P)} \frac{dS'}{\rho\rho'}.
$$

This last integral is known to be bounded.[30]

LEMMA 2a. *$K_1(M, N)$ is bounded uniformly as τ approaches zero.*

We suppose M, N, G on $S(\tau)$ correspond respectively to Q, P, R on S_0. For simplicity of notation we denote $S(\tau)$ by S. Since MN converges uniformly to QP as τ approaches zero, given a positive number δ, then τ_1 exists so that for all $\tau \leq \tau_1$, either $MN \geq 2\delta$ uniformly or $MN \leq 2\delta$ uniformly. Hence, we consider two cases as before.

Case 1. $MN \geq 2\delta > 0$. In this case we write

$$
4\pi^2 K_1(M, N) = 4\pi^2 \int_{S-S(\delta,M)-S(\delta,N)} K(M, G)K(G, N)\,dS_G + 4\pi^2 \int_{S(\delta,M)} + 4\pi^2 \int_{S(\delta,N)}
$$

[30] Kellogg, *Potential Theory*, pp. 302-303.

Denoting the terms on the right by I'_1, I'_2, and I'_3 respectively, we have

$$| I'_1 | \leq \frac{1}{\delta^4} m(S)$$
$$| I'_2 | \leq \frac{1}{\delta^2} \int_{S(\delta,M)} \frac{| \cos \sphericalangle (MG, n_G) |}{\overline{MG}^2} dS_G.$$
$$| I'_3 | \leq \frac{1}{\delta^2} \int_{S(\delta,N)} \frac{| \cos \sphericalangle (GN, n_N) |}{\overline{GN}^2} dS_G.$$

Obviously $| I'_1 |$ is uniformly bounded as τ approaches zero. From Lemma 3 it follows that $| I'_2 |$ and $| I'_3 |$ are uniformly bounded.

Case 2. $MN \leq 2\delta$. Let $\delta' = 4\delta$. Now we have

$$4\pi^2 K_1(M, N) = 4\pi^2 \int_{S-S(\delta',N)} K(M, G) K(G, N) dS_G + 4\pi^2 \int_{S(\delta',N)} .$$

Denoting the terms on the right by J'_1 and J'_2 respectively, we have at once that

$$| J'_1 | \leq \frac{1}{64\delta^4} m(S),$$

and hence $| J'_1 |$ is uniformly bounded. Let $C(\delta', N)$ be the circle with center N and radius δ' and lying in the tangent plane to S at N. Let ρ_I, ρ'_I be the projections of NG and MG respectively in the tangent plane at N, and let dS'' be the projection of dS_G on the tangent plane at N. Without affecting the validity of the proof we may suppose δ small enough so that the inequalities corresponding to (2. 1′) hold uniformly as τ approaches zero. From the inequality (2. 2′) we have evidently

$$| J'_2 | < 80\gamma^2 \int_{C(\delta',N)} \frac{dS''}{\rho_I \rho'_I}.$$

But since dS''/dS', ρ_I/ρ, ρ'_I/ρ' all approach 1 uniformly as τ approaches zero, $\int_{C(\delta',N)} \frac{dS''}{\rho_I \rho'_I}$ approaches $\int_{C(\delta',P)} \frac{dS'}{\rho\rho'}$ uniformly as τ approaches zero. We have seen that $\int_{C(\delta',P)} \frac{dS'}{\rho\rho'}$ is bounded, which completes the proof of the lemma.

3.

LEMMA 3a. *$g_n(M, P)$ is continuous in P for P on S_0 and M in T^+.*

Let M be a fixed point of T^+ with minimum distance D from points of S_0. It is evident that the second term on the right in equation (3. 19) is continuous

in P. For the first term we have

$$\left| \int_{S_0} k(R,P) \frac{\cos \sphericalangle (MR, n_R)}{\overline{MR}^2} dS_R - \int_{S_0} k(R,P_1) \frac{\cos \sphericalangle (MR, n_R)}{\overline{MR}^2} dS_R \right| \leq I + J;$$

where

$$I = \frac{1}{D^2} \int_{S_0(\delta,P)} | k(R,P) - k(R,P_1) | \, dS_R$$

$$J = \frac{1}{D^2} \int_{S_0 - S_0(\delta,P)} | k(R,P) - k(R,P_1) | \, dS_R.$$

We may fix δ so that I is as small as we like. Suppose $\overline{PP_1} < \delta/2$. To complete the proof we have only to show that $k(R,P_1)$ approaches $k(R,P)$ as P_1 approaches P, with $RP > \delta/2$ and $RP_1 > \delta/2$. But with this restriction on R, it is seen that each term in $k(R,P_1)$ as given by (3.17) approaches the corresponding term in $k(R,P)$ as P_1 approaches P. Hence the lemma is true.

4.

We shall make use of the following theorem due to J. Radon:[31]

THEOREM. *Let $f_1, f_2, \cdots$ be a sequence of completely additive functions of point sets whose total variation over J is uniformly bounded and let E_0 be a closed set in J. Then we can pick out a subsequence $\{f_{n^*}\}$ of the functions f_n and a completely additive function of point sets, $f(e)$, so that for each function $\phi(P)$ continuous on E_0*

$$\lim_{n^* \to \infty} \int_{E_0} \phi(P) \, df_{n^*}(e_P) = \int_{E_0} \phi(P) \, df(e_P). \tag{4.1'}$$

It is readily seen that the function $f(e)$ and the sequence f_{n^*} so chosen will also satisfy the relation

$$\lim_{n^* \to \infty} \int_{E_0} \phi_{n^*}(P) \, df_{n^*}(e_P) = \int_{E_0} \phi(P) \, df(e_P) \tag{4.2'}$$

where the functions $\phi_n(P)$ are continuous and converge uniformly to $\phi(P)$. This relation is derived from the inequality

[31] J. Radon, "Theorie und Andwendungen der absolut additiven Mengenfunktionen," *Wiener Akademie Sitzungsberichte*, 122. 2, II (a), (1913), p. 1337.

$$\left| \int_{E_0} \phi_{n^*} df_{n^*} - \int_{E_0} \phi df \right|$$
$$\leq \left| \int_{E_0} \phi_{n^*} df_{n^*} - \int_{E_0} \phi_{n^*} df \right| + \left| \int_{E_0} \phi_{n^*} df - \int_{E_0} \phi df \right| .$$

The first term on the right approaches zero as n^* becomes infinite, by the theorem above. The second term approaches zero as n^* becomes infinite on account of the uniform convergence of ϕ_{n^*} to ϕ. Hence we have the relation (4. 2′). The subsequence $f_{n^*}(e)$ satisfying (4. 1′) is said to converge to the function $f(e)$ in the *weak sense.*

Extension for normal families of surfaces. If $\{f_n(e)\}$ is a sequence of completely additive functions of point sets defined on the subsequence $\{S_n\}$ of a normal family of surfaces and whose total variations are uniformly bounded, and if the sequence $\{S_n\}$ converges to the surface S_0 of the normal family as n becomes infinite, then we may pick out a subsequence f_{n^*} of the f_n's and a completely additive function of point sets $f(e)$ defined on S_0 such that for each $\phi(P_n)$, defined for P_n on S_n and continuous in P_n and converging uniformly to $\phi(P)$ defined on S_0

$$\lim_{n^* \to \infty} \int_{S_n^*} \phi(P_{n^*}) \, df_{n^*}(e_{P_{n^*}}) = \int_{S_0} \phi(P) \, df(e_P).$$

This extension is immediate on account of the 1 : 1 correspondence between points of S_0 and points of S_{n^*} for n^* sufficiently large. We define $f_{n^*}(e)$ on S_0 equal to $f_{n^*}(e')$, where e' on S_{n^*} is the set of points corresponding to e on S_0, and then apply (4. 2′). We say that the functions f_{n^*} on S_{n^*} converge in the *weak sense* to $f(e)$ on S_0.

5.

LEMMA 4a. *The functions $g_n(M, N)$ converge uniformly to $g_n(M, P)$ as τ approaches zero, M being a fixed point in T^+ [where N on $S(\tau)$ corresponds to P on S_0].*

For convenience, we denote $S(\tau)$ by S. Let τ be small enough so that M is interior to S as well as S_0. In the formula for $g_n(M, N)$ corresponding to (3. 19) it is obvious that the second term converges uniformly to the corresponding term in the formula (3. 19) as τ approaches zero. For the first term we have, letting G on S correspond to R on S_0,

$$\text{(5. 1′)} \quad \left| \int_{S_0} k(R, P) \frac{\cos \sphericalangle (MR, n_R)}{\overline{MR}^2} dS_R \right.$$
$$\left. - \int_{S} k(G, N) \frac{\cos \sphericalangle (MG, n_G)}{\overline{MG}^2} dS_G \right| \leq I + J,$$

where

$$I = \left| \int_{S_0(\delta,P)} k(R,P) \frac{\cos \sphericalangle (MR, n_R)}{\overline{MR}^2} dS_R - \int_{S(\delta,P)} k(G,N) \frac{\cos \sphericalangle (MG, n_G)}{\overline{MG}^2} dS_G \right|$$

$$J = \left| \int_{S_0 - S_0(\delta,P)} k(R,P) \frac{\cos \sphericalangle (MR, n_R)}{\overline{MR}^2} dS_R - \int_{S-S(\delta,P)} k(G,N) \frac{\cos \sphericalangle (MG, n_G)}{\overline{MG}^2} dS_G \right| .$$

Given $\epsilon > 0$ we may fix δ so that

$$I < \epsilon/2$$

since both integrals of (5. 1′) converge. Now we let τ_1 be less than δ so that N is in $S(\delta, P)$. The integrand of the second integral of J converges uniformly to that of the first integral as τ approaches zero. Also dS_R/dS_G approaches 1 uniformly as τ approaches zero. Hence τ_1 exists so that

$$J < \epsilon/2$$

for all $\tau \leq \tau_1$. Hence the lemma is proved.

THE RICE INSTITUTE,
HOUSTON, TEXAS.

FORMAL SOLUTIONS OF IRREGULAR LINEAR DIFFERENTIAL EQUATIONS. PART II.

By Frances Thorndike Cope.

In the first part of this paper[1] we proved the fundamental formal existence theorem for the linear differential equation

$$a_0(x)y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \cdots + a_n(x)y(x) = 0, \qquad a_0(x) \not\equiv 0,$$

in which the coefficients $a_i(x)$ are rational functions of x. We shall now prove the converse theorem. In § 6 we shall show the equivalence of the equation (1) to the linear differential system

$$y'_i(x) = \sum_{j=1}^{n} a_{ij}(x)y_j(x) \qquad (i = 1, 2, \cdots, n)$$

in which the $a_{ij}(x)$ are rational functions of x, and hence to the most general linear system with rational coefficients. Sections 7 and 8 are devoted to general theorems on formal reducibility and formal equivalence at ∞ respectively. The definitions and general notation of the first part will be retained.

5. Converse of fundamental theorem. The converse of Theorem I may be stated as follows:

Theorem VI. *Any set of n linearly independent formal series $y_1, y_2, \cdots, y_n$ of the general type* (2) *which has the properties,*

i) *that if one member of the set is of anormal form then all possible determinations of this expression are linearly dependent on $y_1, y_2, \cdots, y_n$,*

ii) *that the complete set consists of one or more subsets of the form* (4), *determines an essentially unique equation of the form* (1), *of the n-th order, for which these series are a complete set of formal solutions. It is, in fact, the equation*

$$| y_1, y'_2, \cdots, y_n^{(n-1)}, y^{(n)} | = 0. \tag{40}$$

This equation is obviously satisfied by each of the given series, since on replacing y by y_i we have two columns of the determinant identical. It is actually of n-th order, for the coefficient of $y^{(n)}$ is the Wronskian of the given

[1] This appeared in *American Journal of Mathematics*, vol. 56 (1934), pp. 411-437.

set of solutions, which is not zero since the set is linearly independent. Moreover the equation is unique, for the equations

$$a_0 y_i^{(n)} + a_1 y_i^{(n-1)} + \cdots + a_n y_i = 0 \qquad (i = 1, 2, \cdots, n)$$

constitute a system of n linear equations in the n ratios

$$a_1/a_0,\ a_2/a_0, \cdots, a_n/a_0,$$

in which the determinant of the coefficients is not zero, since it is precisely the Wronskian, $\Delta(y_1, \cdots, y_n)$, of the set of solutions, which are assumed linearly independent. It remains to show that the equation is actually of the type (1), that is, that the coefficients are simple formal series.

To prove this we observe first that any exponential factor which occurs in any solution will also occur in all of its derivatives, and hence in all the elements of one column of the determinant, so that it can be factored out from the equation. Thus the equation can be freed of exponential factors.

We have next to consider the possibility that it may involve $\log x$. Let $y_1, \cdots, y_m$ be a set of solutions of the form (4). Then by successive differentiations we find that

$$y_{k+1}^{(i)}(x) = A_{0i}(x) \log^k x + \cdots + \frac{k!}{j!\,(k-j)!} A_{ji}(x) \log^{k-j} x + \cdots + A_{ki}(x)$$
$$(k = 0, 1, \cdots, m-1),$$

where the coefficients

$$A_{j0}(x) = s_j(x), \qquad A_{ji}(x) = A'_{j,i-1}(x) + x^{-1} A_{j-1,i-1}(x) \qquad (0 < i),$$

are independent of k. The first m columns of the determinant then are

$$A_{00},\ A_{00}\log x + A_{10}, \cdots, A_{00}\log^{m-1} x + \cdots + A_{j0}\frac{(m-1)\cdots(m-j)}{j!}\log^{m-j-1} x + \cdots + A_{m-1,0},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$A_{0n},\ A_{0n}\log x + A_{1n}, \cdots, A_{0n}\log^{m-1} x + \cdots + A_{jn}\frac{(m-1)\cdots(m-j)}{j!}\log^{m-j-1} x + \cdots + A_{m-1,n},$$

and by subtracting $\log^{k-1} x$ times the first column from the k-th column $(2 \leq k \leq m)$ the leading term of each element is eliminated from all columns 2 to m, i. e. the highest power of $\log x$ is reduced by one. Then by subtracting $(k-1)\log^{k-2} x$ times the second column from the k-th column $(3 \leq k \leq m)$ the highest power of $\log x$ which appears is again reduced, and by continuing

the process we can eliminate step by step all the terms involving $\log x$ in these columns of the determinant. Since this set of solutions is typical of those in which $\log x$ occurs, the expansion of the determinant is evidently free from $\log x$. Thus the equation (40), since it does not involve $\log x$ or any exponential factor, is of the required form, that is, has coefficients which are formal series in descending powers of $x^{1/mp}$.

Since any determination of a series $y_i(x)$ of anormal form is linearly dependent on $y_1, y_2, \cdots, y_n$, it must also be a solution of the equation (40). Consequently the equation obtained is essentially independent of the choice of $\arg x^{1/mp}$, $\arg x^{1/p}$ being assumed fixed. Its coefficients therefore do not involve fractional powers of $x^{1/p}$ but are expressible as formal series in $x^{1/p}$.

6. Reduction of general linear system to a single equation. It is well known that any system of linear differential equations can be reduced to a system of linear equations of the first order, homogeneous if the original system is homogeneous. For example, the homogeneous equation of the n-th order in one variable, equation (1), can always be reduced to a system of n linear homogeneous equations of the first order in n variables, in particular by taking

$$y_1(x) = y(x), \qquad y_i(x) = y'_{i-1}(x) \qquad (i = 2, 3, \cdots, n),$$

in which case the linear system is

$$\begin{aligned} y'_i(x) &= y_{i+1}(x) \\ a_0(x)y'_n(x) &= -a_1(x)y_n(x) - \cdots - a_n(x)y_1(x) \end{aligned} \qquad (i = 1, 2, \cdots, n-1)$$

and has rational coefficients if the coefficients of the equation (1) are rational. It is also true, though less familiar, that any linear system

$$(41) \qquad y'_i(x) = \sum_{j=1}^{n} a_{ij}(x) y_j(x) \qquad (i = 1, 2, \cdots, n)$$

with rational coefficients can be reduced to a single equation (1) of the n-th order with rational coefficients.

To prove this we let

$$y(x) = \lambda_1(x)y_1(x) + \lambda_2(x)y_2(x) + \cdots + \lambda_n(x)y_n(x)$$

where the functions $\lambda_i(x)$ are for the present arbitrary rational functions. Then on differentiating repeatedly and substituting each time the values of $y'_i(x)$ from the equations (41) we have $n+1$ equations

$$y^{(i-1)}(x) = \sum_{j=1}^{n} \lambda_{ij}(x) y_j(x) \qquad (i = 1, 2, \cdots, n+1), \tag{42}$$

in which the right-hand terms are linear combinations of $y_1(x), \cdots, y_n(x)$. Consequently there must be at least one linear homogeneous equation between the left-hand terms, that is, an equation of the form

$$a_0(x) y^{(n)}(x) + \cdots + a_n(x) y(x) = 0$$

in which at least one of the coefficients is different from zero.

In fact we can choose the $\lambda_i(x)$ so that $a_0(x) \not\equiv 0$, since they can be so chosen that the determinant of the coefficients of $y_1, y_2, \cdots, y_n$ in the first n of the above $n+1$ equations is not identically zero. The elements of the i-th row of this determinant are

$$\gamma_{ij} = \lambda_j^{(i-1)} + P_{ij} \qquad (j = 1, 2, \cdots, n),$$

where P_{ij} is a polynomial in $\lambda_1, \cdots, \lambda_n$, their first $i-2$ derivatives, a_{rs} $(r, s = 1, \cdots, n)$, and their derivatives. Hence at any point $x = x_0$ which is not a pole of one of the coefficients $a_{rs}(x)$, the elements of the determinant can be given arbitrary values by assigning suitable values first to $\lambda_1(x_0), \cdots, \lambda_n(x_0)$, then to $\lambda'_1(x_0), \cdots, \lambda'_n(x_0)$, and to the derivatives of higher orders successively, up to $\lambda_1^{(n-1)}(x_0), \cdots, \lambda_n^{(n-1)}(x_0)$. It is always possible to determine a rational function $\lambda_j(x)$ such that $\lambda_j^{(i)}(x_0) = c_{ij}$, where the c_{ij} $(i = 0, 1, \cdots, n)$ are arbitrary constants. In fact, the polynomial

$$\lambda_j(x) \equiv c_{0j} + c_{1j}(x - x_0) + \cdots + \frac{c_{ij}}{i!}(x - x_0)^i + \cdots + \frac{c_{n-1,j}}{(n-1)!}(x - x_0)^{n-1}$$

is such a function. Thus it is possible to determine $\lambda_1(x), \cdots, \lambda_n(x)$ so that $|\gamma_{ij}(x_0)| \neq 0$, and hence $a_0(x) \not\equiv 0$, and we have an equation of the n-th order in $y(x)$. Moreover the coefficients in this equation are rational combinations of the coefficients of the linear system and therefore are rational functions of x.

The solution of the linear system in terms of $y(x)$ is then obtained by solving simultaneously the first n equations of the set (42) for $y_1, \cdots, y_n$. Since $|\gamma_{ij}|$, the determinant of the coefficients, has been made different from zero, Cramer's rule may be applied, and the solution will be of the form

$$y_i(x) = \beta_{i1} y^{(n-1)}(x) + \cdots + \beta_{in} y(x) \qquad (i = 1, 2, \cdots, n) \tag{43}$$

where the $\beta_{i1}, \cdots, \beta_{in}$ are rational functions of x. Thus each solution of the n-th order equation in $y(x)$ will lead to a solution of the linear system, and we have

THEOREM VII. *Any system of homogeneous lin of the first order with rational coefficients can be re homogeneous differential equation in one variable wi*

We may therefore regard either the single homo order in one variable or the system of n homogenec order in n variables as essentially equivalent to the n linear system. Since, however, the system of n equa often the more convenient to use, we shall derive t Theorem I for such a system.

From the equations (43) which give the soluti (41) in terms of the solutions of a single equation (each subset,

$$y^{[m+1]}(x) = s_0(x)\log^m x + \cdots + \frac{m!}{j!(m-j)!} s_j(x)$$

of solutions of the single equation there corresponds

$$y_i^{[m+1]}(x) = \alpha_{i0}(x)\log^m x + \cdots + \frac{m!}{j!(m-j)!} \alpha_{ij}(x)$$
$$(i = 1, 2, \cdots,$$

of solutions of the corresponding system.

If $y^{[m+1]}(x)$ is of the form (2) we find that its i

$$y^{[m+1](i)}(x) = A_{0i}(x)\log^m x + \cdots + \frac{m!}{j!(m-j)!} A_{ji}(x)\log^{m-j} x$$

where $A_{ji}(x)$ is defined, for $(j = 0, 1, \cdots, m)$, by

$$A_{j0}(x) = s_j(x), \quad A_{ji}(x) = A'_{j,i-1}(x) + x^{-1}A_{j-1,i-1}(x$$

Consequently, after substituting these values in the collect coefficients of powers of $\log x$ and have, for $(m = 0, 1, \cdots, k)$,

$$y_i^{[m+1]}(x) = \sum_{j=0}^{m} \frac{m!}{j!(m-j)!} [\beta_{i1}(x)A_{j,n-1}(x) + \cdots$$

which we may write in the form

$$y_i^{[m+1]}(x) = \sum_{j=0}^{m} \frac{m!}{j!(m-j)!} \alpha_{ij}(x)\log^{m-j} x \quad (i = 1, 2, \cdot$$

where the coefficients $\alpha_{ij}(x)$ do not depend on m, that is, in a form similar to that of $y^{[m+1]}(x)$. Hence each such set of solutions, $y^{[m+1]}(x)$ $(m=0,1,\cdots,k)$, leads to a corresponding set of solutions of similar form, $y_1^{[m+1]}(x),\cdots,y_n^{[m+1]}(x)$ $(m=0,1,\cdots,k)$.

From the equation by which $y(x)$ is introduced it is clear that if a set of solutions of the linear system (41) is linearly dependent then so is the corresponding set of solutions of the equation (1), for if there exist constants $c_1,\cdots,c_n$, not all zero, such that

$$c_1y_i^{[1]}+c_2y_i^{[2]}+\cdots+c_ny_i^{[n]}=0 \qquad (i=1,2,\cdots,n),$$

then

$$c_1y^{[1]}+\cdots+c_ny^{[n]}=\sum_{i=1}^{n}\lambda_i[c_1y_i^{[1]}+\cdots+c_ny_i^{[n]}]=0$$

also. Therefore as a corollary of Theorem I we have

THEOREM VIII. *Any linear system (41) in which the coefficients are formal series in descending powers of $x^{1/p}$, p being a positive integer, has always a complete set of n distinct formal solutions of the general type*

$$y_i(x)=s_{i0}(x)\log^k x+s_{i1}(x)\log^{k-1}x+\cdots+s_{ik}(x) \quad (i=1,2,\cdots,n),$$

and the complete set consists of subsets of the form

$$y_i^{[j+1]}(x)=\sum_{l=0}^{j}\frac{j\,!}{l\,!(j-l)\,!}\,s_{il}(x)\log^{j-l}x \quad (i=1,2,\cdots,n;\ j=0,1,\cdots,k).$$

From the complete set of solutions for the homogeneous equation (1) a particular solution of the corresponding non-homogeneous equation

$$a_0(x)y^{(n)}(x)+\cdots+a_n(x)y(x)=r(x), \tag{44}$$

where $r(x)$ is a formal series in descending powers of $x^{1/p}$, can be found by the method of variation of parameters. This solution is

$$Y(x)=c_1(x)y_1(x)+\cdots+c_n(x)y_n(x),$$

where

$$c'_i(x)=r(x)\Delta_{ni}/\Delta\ ^{2}$$

and Δ_{ni} denotes the cofactor of the element in the n-th row and i-th column of $\Delta(y_1,y_2,\cdots,y_n)$.

Let the exponential factors of $y_1,y_2,\cdots,y_n$ be denoted by

---

[2] By the process used to show that equation (40) involves neither $\log x$ nor exponential factors, Δ can be shown to be of the form (7).

$$e^{Q_1(x)}, e^{Q_2(x)}, \cdots, e^{Q_n(x)}$$

respectively. Then the quotient Δ_{ni}/Δ has the exponential factor $e^{-Q_i(x)}$. From the proof of the lemma on page 26 it is clear that the solution of the non-homogeneous equation of first order has the same exponential factor as the right-hand term of the equation. Hence $c_i(x)$ has the exponential factor $e^{-Q_i(x)}$, and the products $c_i(x)y_i(x)$ $(i=1,2,\cdots,n)$ have no exponential factors. The particular solution $Y(x)$ of the non-homogeneous equation of the n-th order is therefore an *elementary* formal solution and in fact one in which no exponential factor occurs, and we have the

THEOREM IX. *Any non-homogeneous linear differential equation* (44) *in which both the coefficients and the right-hand terms are formal series in descending powers of* $x^{1/p}$, *has at least one particular formal solution which is of the elementary type* (2) *and in fact does not involve an exponential factor.*

The same general method may be employed to determine a particular solution of the non-homogeneous linear system

$$y'_i(x) = \sum_{j=1}^{n} a_{ij}(x)y_j(x) + r_i(x) \qquad (i=1,2,\cdots,n)$$

in which the $r_i(x)$ as well as the coefficients $a_{ij}(x)$ are formal series, and an elementary solution containing no exponential factor is obtained in this case also.

7. Reducibility. In the course of the proof of Theorem I we have seen that an equation of the type (1′) and of order greater than one is always reducible if the basic integer p is suitably chosen. For an arbitrarily chosen admissible basic integer, however, the equation is not necessarily reducible. The reducibility of the equation is intimately related to the character of its complete set of formal solutions. In order to show the relationship we introduce the concept of a *natural family* of solutions.

Let $y_1(x),\cdots,y_n(x)$ be a set of n linearly independent solutions of the equation (1′). Then any linear combination of them with constant coefficients is also a solution of this equation. The aggregate of such combinations which are of the form (2) constitutes a family of solutions, which has the property that if one determination of multiple-valued solution belongs to it then so do the other determinations. Any set of elementary solutions of the equation which has these two properties, namely,

i) *that any member of the set can be expressed as a linear combination of m linearly independent members of the set,*

ii) *that if one determination of a multiple-valued solution belongs to the set then so do its other determinations*

is called a *natural family* of solutions, and the number m of linearly independent members is called the *order* of the family. For example, a single non-logarithmic solution of normal form constitutes a natural family. This definition of natural family clearly, like those of reducibility and of normal and anormal series, is relative to a particular basic integer p.

It is not difficult to show that among the various sets of m linearly independent members, by which a given natural family can be generated, there is at least one of the type described in Theorem I and required by the hypotheses of Theorem VI. Hence we have the

THEOREM VI′. *Any natural family of solutions determines an equation of type* (1′), *of order equal to the order of this family, which has the members of this natural family, and only these, as its formal solutions.*

If we add to any given natural family, say F_1, of order n_1, one or more new members such that the set thus formed is also a natural family, say F_2, of order n_2 $(> n_1)$, the new family F_2 will determine an equation of order n_2. If the two corresponding equations are denoted by $L_1(y) = \theta$ and $M(y) = 0$ respectively, then it is clear, since any solution of the first is also a solution of the second, that the second may be expressed as $L_2(L_1(y)) = 0$, that is, that L_1 is a symbolic factor of M. Similarly, to any expanding sequence of natural families $F_1, F_2, \cdots, F_m$, there corresponds a sequence of equations,

$$L_1(y) = 0,\ L_2(L_1(y)) = 0, \cdots, L_m(L_{m-1} \cdots (L_1(y))) = 0,$$

such that the formal solutions of the i-th equation are precisely the members of F_i $(i = 1, 2, \cdots, m)$.

On the other hand, if the differential expression $L(y)$ can be factored symbolically, that is, if $L(y) = M(N(y))$, where M and N are differential operators of the same type as L, then the solutions of the equation $N(y) = 0$ are solutions of the equation $L(y) = 0$ also, and hence form a sub-family of the natural family of solutions of the latter equation. Consequently any factorization $L \equiv L_m L_{m-1} \cdots L_1$ of the differential operator $L(y)$ determines an expanding sequence of natural families $F_1, F_2, \cdots, F_m$ such that F_i consists of the formal solutions of the equation

$$L_i(L_{i-1} \cdots (L_1(y))) = 0 \qquad (i = 1, 2, \cdots, m).$$

Furthermore, if the factors $L_1, \cdots, L_m$ are irreducible, then the corresponding sequence of natural families is such that there is no intermediate natural family distinct from those of the sequence, and conversely.

This result is precisely analogous, in statement and proof, to the theorem on reducibility stated by Birkhoff for the linear difference equations.[3] It may be stated as follows:

THEOREM X. *To any decomposition of $L(y)$, of order n, into irreducible symbolic factors $L_1, L_2, \cdots, L_m$ such that $L \equiv L_m L_{m-1} \cdots L_1$, there corresponds a sequence of natural families $F_1, F_2, \cdots, F_m$, each containing the preceding as a sub-family, such that there exist no intermediate natural families, and such that the general solution of $L_1(y) = 0$ is furnished by F_1, of $L_1 L_2(y) = 0$ by F_2, etc.*

Conversely, to any set of natural families $F_1, F_2, \cdots, F_m$ (of formal solutions of $L(y) = 0$), each containing the preceding as a sub-family, but such that there exist no intermediate natural families, there corresponds an irreducible factorization of $L(y)$, which is essentially unique.

8. **Equivalence.** For certain purposes it is convenient to focus attention on the set of *distinct* natural families into which a given natural family can be divided, rather than on an expanding sequence of natural families contained in it. It is clear that each set of solutions (4), with their various determinations if they are multiple-valued, determines a distinct natural family of solutions of the equation (1). Similarly the complete set of solutions of the linear system

$$y'_i(x) = \sum_{j=0}^{n} a_{ij}(x) y_j(x) \qquad (i = 1, 2, \cdots, n) \tag{45}$$

consists of distinct sets of the form

$$y^{[\lambda]}_{i1}(x), y^{[\lambda]}_{i2}(x), \cdots, y^{[\lambda]}_{i,k+1}(x) \quad (\lambda = 1, 2, \cdots, m;\ i = 1, 2, \cdots, n), \tag{46}$$

where $y^{[1]}_{ij}(x), \cdots, y^{[m]}_{ij}$ denote the m distinct determinations of the solution $y_{ij}(x) = s_{i0}(x) \log^{j-1} x + \cdots + s_{i,j+1}(x)$ for $y_i(x)$.

This decomposition of the fundamental set of solutions into distinct sets which correspond to distinct natural families is useful in studying the equivalence of linear systems. In matrix notation the system (45) is written as $Y'(x) = A(x) Y(x)$ where $A(x)$ denotes the matrix $\| a_{ij}(x) \|$, and the matrix solution $S(x)$ is the n-rowed square matrix formed by the fundamental set

[3] *Loc. cit.*, pp. 238-241.

of solutions. (Hence $|S(x)| \neq 0$. A linear transformation $Y(x) = B(x)\bar{Y}(x)$ ($|B(x)| \neq 0$), in which the elements of $B(x)$ are series of the form $b(x) = b_0 + b_1 x^{-1/mp} + b_2 x^{-2/mp} + \cdots$ then takes the original equation (45) into another

$$\bar{Y}'(x) = \bar{A}(x)\bar{Y}(x) \qquad (\bar{A}(x) = B^{-1}(x)[A(x)B(x) - B'(x)]),$$

of the same type, and the two equations are said to be *formally equivalent at* ∞. If all the elements of $B(x)$ are of normal form the equations are called *properly equivalent*; otherwise they are called *improperly equivalent*.[4] If $S(x) = B(x)E(x)$ ($|B(x)| \neq 0$), so that $E(x)$ is a solution of the new equation, then we have $\bar{A}(x) = E'(x)E^{-1}(x)$.

From the form of the sets (46) we find that such a factorization of $S(x)$ is always possible. Let the first $m(k+1)$ columns of $S(x)$ consist of a set of solutions (46), with the common exponential factor $e^{Q(x)}$; let the highest positive power of x which occurs in any of the coefficients $s_{ij}(x)$ be the r-th;[5] let $e_{1\lambda}$ denote the λ-*th* determination of $x^r e^{Q(x)}$; and let $s_{ij}^{[\lambda]}(x)$ be the λ-th determination of $s_{ij}(x)/e_{1\lambda}$. Then $s_{ij}^{[\lambda]}(x)$ is a simple formal series with no terms in positive powers of $x^{1/mp}$, and these columns of $S(x)$ can evidently be obtained in a product $B(x)E(x)$ if the first $m(k+1)$ columns of $B(x)$ are

$$\underbrace{\begin{matrix} \underbrace{\begin{matrix} s_{10}^{[1]}, & \cdots, & s_{1k}^{[1]}; \\ \cdot & \cdot & \cdot \\ s_{n0}^{[1]}, & \cdots, & s_{nk}^{[1]}; \end{matrix}}_{k+1 \text{ columns}} & \begin{matrix} \cdots; \\ \\ \cdots; \end{matrix} & \underbrace{\begin{matrix} s_{10}^{[m]}, & \cdots, & s_{1k}^{[m]} \\ \cdot & \cdot & \cdot \\ s_{n0}^{[m]}, & \cdots, & s_{nk}^{[m]} \end{matrix}} \end{matrix}}_{m \text{ sets of } k+1 \text{ columns each}}$$

and the first $m(k+1)$ rows of $E(x)$ are

$$\begin{matrix}
e_{11}, e_{11}\log x, \cdots, e_{11}\log^k x; & 0, \cdots\cdots\cdots\cdots\cdots\cdots, 0 \\
0, \; e_{11}, \cdots, k e_{11}\log^{k-1} x; & 0, \cdots\cdots\cdots\cdots\cdots\cdots, 0 \\
\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\
0, \cdots\cdots, e_{11}; & 0, \cdots\cdots\cdots\cdots\cdots\cdots \\
\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\
0, \cdots\cdots\cdots\cdots\cdots, 0; & e_{1m}, \cdots, e_{1m}\log^k x; \; 0, \cdots, 0 \\
\cdots\cdots\cdots\cdots & \cdots\cdots\cdots\cdots\cdots\cdots \\
0, \cdots\cdots\cdots\cdots\cdots, 0; & 0, \cdots, e_{1m}; \quad 0, \cdots, 0.
\end{matrix}$$

Similarly each set of solutions (46), i. e., each distinct natural family of

[4] Cf. Birkhoff, *loc. cit.*, p. 242.

[5] If no positive powers occur let $r = 0$.

solutions, determines a set of columns of $B(x)$ and of rows of $E(x)$, and we have

THEOREM XI. *An arbitrary linear system*

$$Y'(x) = A(x)Y(x), \tag{45'}$$

in which the elements of the matrix $A(x)$ are formal series in descending powers of $x^{1/p}$, is equivalent (improperly) to a normal system

$$\bar{Y}'(x) = E'(x)E^{-1}(x)\bar{Y}(x),$$

in which the matrix $E(x)$ consists of zeros except for blocks, along the principal diagonal, of the form

$$\begin{matrix} e_{\nu\lambda}, & e_{\nu\lambda}\log x, \cdots, & e_{\nu\lambda}\log^{k_\nu} x \\ 0, & e_{\nu\lambda}, \quad \cdots, & k_\nu e_{\nu\lambda}\log^{k_\nu - 1} x \\ \cdot & \cdot \quad \cdot \quad \cdot & \cdot \\ 0, & \cdots\cdots\cdots, & e_{\nu\lambda}, \end{matrix}$$

where $e_{\nu\lambda}$ denotes the λ-th determination of the factor $x^{r}e^{Q_\nu(x)}$ in the ν-th set of solutions (46) of the system (45').

To illustrate the application of this result consider the special case in which the formal solutions are all non-logarithmic and of normal type. Then there are n distinct natural families, the matrix $E(x)$ consists entirely of zeros except for the elements $x^r e^{Q_1(x)}, \cdots, x^r e^{Q_n(x)}$ on the principal diagonal, and the normal form is

$$y'_i(x) = [Q'_i(x) + rx^{-1}]y_i(x) \qquad (i = 1, 2, \cdots, n).$$

RADCLIFFE COLLEGE.

ON THE ALGEBRAIC PROBLEM CONCERNING THE NORMAL FORMS OF LINEAR DYNAMICAL SYSTEMS.

By JOHN WILLIAMSON.

Introduction. Let m be the number of degrees of freedom of a linear conservative dynamical system and let the point $(q_1, q_2, \cdots, q_m, p_1, p_2, \cdots, p_m)$ of the phase space be denoted by $x = (x_1, x_2, \cdots, x_{2m})$. A system of $2m$ ordinary differential equations of the first order, which are homogeneous, linear and do not contain t explicitly, is a canonical system if, and only if, there exists a symmetric matrix A, such that the differential equations may be written in the form

$$\text{(i)} \qquad Bdx/dt = Ax,$$

where B is the skew symmetric matrix $\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}$, and E the unit matrix of order m. In fact, apart from a factor 2, A is simply the matrix of the $2m$-ary quadratic form, which represents the Hamiltonian function. A non-singular matrix P is said to be a *Hamiltonian* matrix, if the transformation $x = Py$ sends every differential system of the form (i) into a differential system of the same form.

It has been pointed out by Wintner [1] that, if the system (i) is transformed into the system

$$\text{(ii)} \qquad Bdy/dt = Cy$$

by the transformation $x = Py$, then P is a Hamiltonian matrix if and only if

$$\text{(iii)} \qquad P'BP = sB \quad \text{and} \quad P'AP = sC.$$

In the following pages we use this result to determine a normal form for the system of equations (i). Equations (iii) imply,

$$P'(A - \lambda B)P = s(C - \lambda B),$$

or

$$P'_1(A - \lambda B)P_1 = \pm (C - \lambda B),$$

where $P_1 = (1/\sqrt{|s|})P$. Since $\begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix} B \begin{pmatrix} 0 & E \\ E & 0 \end{pmatrix} = -B$, we have either

[1] A. Wintner, "On the linear conservative dynamical systems," *Annali di matematica pura ed applicata*, ser. 4, tomo 13 (1934-1935).

$$P'_1(A-\lambda B)P_1 = C-\lambda B \quad \text{or} \quad P'_2(A-\lambda$$

where $P_2 = P_1\begin{pmatrix}0 & E\\ E & 0\end{pmatrix}$ and $C_2 = \begin{pmatrix}0 & E\\ E & 0\end{pmatrix}C\begin{pmatrix}0 & E\\ E & 0\end{pmatrix}$.

a matrix of matrices of order m, so that $C = \begin{pmatrix}C_{11} & C_{1}\\ C_{21} & C_{2}\end{pmatrix}$

It is therefore apparent that a normal form for the s

from a suitable canonical form of the matrix pencil

Accordingly we first consider the purely algebra

a canonical form for a pencil $A-\lambda B$ where A is

symmetric and non-singular. For the sake of gene

elements of the matrices A and B lie in a commuta

transformation matrices are restricted to have ele

Later we particularize B to be the matrix $\begin{pmatrix}0 & -\\ E & \end{pmatrix}$

applicable to the original dynamical problem. In th

fications, which arise when K is the field of all real

and a list of the possible normal forms of the matrix

case of two degrees of freedom ($m = 2$). These no

elementary divisors of $A-\lambda B$ which may be *real, co*

It is interesting to note that, if any of the elementa

pure imaginary, the elementary divisors alone are n

the normal form.

I. Let A and B be two square matrices of or

commutative field K of characteristic zero. Further

B be skew-symmetric and non-singular, so that

$|B| \neq 0$. If M is any matrix with elements in K,

there exists a non-singular matrix P, with ele

$P^{-1}AB^{-1}P = M$. Hence, if λ is any indeterminate, $P^{-1}($

Accordingly,

$$P'(B^{-1})'(A-\lambda B)B^{-1}P = P'(B^{-1})'P(M-\lambda I$$

where $R = P'(B^{-1})'P$. If $C = B^{-1}P$, we may wr

the form

$$(1) \qquad C'(A-\lambda B)C = R(M-\lambda E$$

As a consequence of (1),

$$R = C'BC \quad \text{and} \quad RM = C'A$$

[2] Cf. C. Lanczos, "Eine neue Transformationstheorie li gen," *Annalen der Physik*, 5 Folge; ser. 653, Band 20 (193

so that R is skew symmetric and RM is symmetric. Therefore

$$RM = (RM)' = M'R' = -M'R. \tag{2}$$

The pencil $A - \lambda B$ is *equivalent* under a non-singular *congruent transformation* with elements in K to the pencil $RM - \lambda R$ and we may, without any risk of confusion, simply say that the two pencils are equivalent. We shall proceed to determine a canonical form for the pencil $A - \lambda B$ by choosing a suitable form for the matrix M and by reducing the matrix R. We first notice that, if W is a non-singular matrix with elements in K and if

$$W'R(M - \lambda E)W = S(M - \lambda E), \tag{3}$$

then

$$S = W'RW \tag{4}$$

and

$$W'RMW = SM = W'RWM \text{ by } (4),$$

so that, since $W'R$ is non-singular,

$$MW = WM. \tag{5}$$

Hence, in the reduction of the matrix R, we are only at liberty to use transformations, whose matrices are commutative with M.

Further, if Q is a non-singular matrix satisfying the equation

$$QM = -M'Q, \tag{6}$$

it follows easily from (2) that

$$R = QG, \tag{7}$$

where

$$GM = MG. \tag{8}$$

If M is a *diagonal block* matrix,

$$M = [M_1, M_2, \cdots, M_t], \tag{9}$$

where M_i is a square matrix of order e_i, we may write G as a matrix of matrices,

$$G = (G_{ij}) \qquad (i, j = 1, 2, \cdots, t),$$

where G_{ij} is a matrix of e_i rows and e_j columns.

If Q_i is a non-singular matrix of order e_i, such that

$$Q_iM_i = -M'_iQ_i \qquad (i = 1, 2, \cdots, t),$$

then the diagonal block matrix

$$Q = [Q_1, Q_2, \cdots, Q_t] \tag{10}$$

is non-singular and satisfies (6).

We now prove,

LEMMA 1. *If* $Q'_i = \rho_i Q_i$ $(i = 1, 2, \cdots, t)$, *i* *non-singular, there exists a non-singular matrix* W *that* $W'QGW = QH$, *where* $H = (H_{ij})$ $(i, j = 1,$ $H_{1j} = H_{j1} = 0$ $(j \neq 1)$.

Proof. Let

$$W = \begin{bmatrix} E_1 & -G_{11}^{-1}G_{12} & -G_{11}^{-1}G_{13} & \cdots \\ 0 & E_2 & 0 & \cdot\ \cdot \\ 0 & 0 & E_3 & \cdot\ \cdot \\ \cdot & \cdot & \cdot & \cdot\ \cdot \\ 0 & 0 & 0 & \cdot\ \cdot \end{bmatrix}$$

where E_i is the unit matrix of order e_i.

Since the matrix G satisfies (8), $G_{ij}M_j =$ $-G_{11}^{-1}G_{1j}M_j = -G_{11}^{-1}M_1G_{1j} = -M_1G_{11}^{-1}G_{1j}$.] mutative with M.

Since R is skew symmetric, it follows from (7)

$$Q_iG_{ij} = -G'_{ji}Q'_j = -\rho_j G' \tag{11}$$

The element in the j-th row, $j \neq 1$, of the first cc

$$-G'_{1j}(G_{11}^{-1})'Q_1.$$

But

$$-G'_{1j}(G_{11}^{-1})'Q_1 = -\rho_1 G'_{1j}Q_1G_{11}^{-1} = -\rho_1^2 Q_j($$

by (11) and the definition of ρ_1. Therefore $W'Q =$ obtained from W' by substituting $-G_{j1}G_{11}^{-1}$ for — in the j-th place of the first column, $j \neq 1$. Sin row and the first column of the product LGW is G in the first row or column are zero, the lemma is pr

If the diagonal block matrix M in (9) is such mutative with M is also a diagonal block matrix $G =$ G_i is a square matrix of order e_i, then

$$G_iM_i = M_iG_i \tag{12}$$

Further, as a consequence of (7), R is a diagonal block matrix $[R_1, R_2, \cdots, R_t]$, where

$$(13) \qquad R_i = Q_i G_i \qquad (i = 1, 2, \cdots, t),$$

and since W is commutative with M, $W = [W_1, W_2, \cdots, W_t]$ and the matrix S defined by (4) is a diagonal block matrix $[S_1, S_2, \cdots, S_t]$, where

$$(14) \qquad S_i = W'_i R_i W_i \qquad (i = 1, 2, \cdots, t).$$

But, apart from the suffixes i, equations (12), (13) and (14) are the same as (8), (7) and (4) respectively. *Therefore, in reducing R to S, we need only consider the reduction of each block R_i separately by matrices commutative with M_i.*

2. *Form of the matrix M.* Since the elements of the matrices A and B lie in the field K, the invariant factors of the pencil $A - \lambda B$ are polynomials in λ with coefficients in K. We shall call the powers of the distinct irreducible factors of the invariant factors, *the elementary factors* (with respect to K) of the pencil. Since A is symmetric and B is skew symmetric, the invariant factors are unaltered, except perhaps in sign, by the interchange of λ and $-\lambda$. Hence each invariant factor is the product of an even polynomial in λ by a power of λ. Accordingly the elementary factors of the pencil $A - \lambda B$ are of three types:

Type a. $[p(\lambda)]^k$ together with $[p(-\lambda)]^k$, where $p(\lambda)$ is an irreducible polynomial but is not an even polynomial in λ and $p(\lambda) \neq \lambda$.

Type b. $[h(\lambda)]^k$ where $h(\lambda) = p(\lambda^2)$ is an even irreducible polynomial in λ.

Type c. λ^k.

We now proceed to determine matrices with elementary factors of types (a), (b) and (c) respectively.[3]

Type (a). Let $p(\lambda)$ be of degree m and let p be any matrix of order m with elements in K, whose characteristic equation is $p(\lambda)$ [4] $= 0$, and let e be the unit matrix of order m. Then, if

$$(15) \qquad \pi = \begin{pmatrix} p & 0 \\ 0 & -p \end{pmatrix}, \qquad \phi = \begin{pmatrix} e & 0 \\ 0 & -e \end{pmatrix}, \qquad \epsilon = \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix},$$

[3] If D is a matrix with elements in K we shall mean by the elementary factors of D, the elementary factors of the pencil $D - \lambda E$, where E is the unit matrix.

[4] We may take as the matrix p the companion matrix of $p(\lambda)$. Cf. J. Williamson, "The equivalence of non-singular pencils of hermitian matrices in an arbitrary field," *American Journal of Mathematics*, vol. 57 (1935), p. 475.

the matrix

(16)
$$N = \begin{pmatrix} \pi & \phi & 0 & \cdot & 0 \\ 0 & \pi & \phi & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \phi \\ 0 & 0 & 0 & \cdot & \pi \end{pmatrix}$$

of order k, considered as a matrix of matrices of order $2m$, has the elementary factors $[p(\lambda)]^k$, $[p(-\lambda)]^k$. For $[p(N)]^k[p(-N)]^k = 0$ and N satisfies no equation of lower degree. We now write (16) in the more convenient form

(17)
$$N = \pi E + \phi U,$$

where

$$E = \begin{pmatrix} \epsilon & 0 & \cdot & 0 \\ 0 & \epsilon & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \epsilon \end{pmatrix} \quad \text{and} \quad U = \begin{pmatrix} 0 & \epsilon & 0 & \cdot & 0 \\ 0 & 0 & \epsilon & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \epsilon \\ 0 & 0 & 0 & \cdot & 0 \end{pmatrix},$$

and proceed to determine a non-singular matrix V satisfying

(18)
$$VN = -N'V.$$

If

(19)
$$T = \begin{pmatrix} 0 & 0 & \cdot & 0 & \epsilon \\ 0 & 0 & \cdot & \epsilon & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \epsilon & \cdot & 0 & 0 \\ \epsilon & 0 & \cdot & 0 & 0 \end{pmatrix},$$

we see immediately that $T^2 = E$ and that

(20)
$$TU = U'T.$$

Further, we can determine a non-singular symmetric matrix q such that [5]

(21)
$$qp = p'q.$$

Since the matrix τ, defined by

(22)
$$\tau = \begin{pmatrix} 0 & -e \\ e & 0 \end{pmatrix},$$

satisfies the equation

(23)
$$\tau\phi = -\phi\tau = -\phi'\tau,$$

it follows that the matrix

(24)
$$V = q\tau T$$

[5] J. Williamson, *loc. cit.*, p. 490.

is a non-singular skew symmetric matrix satisfying (18). In fact,

$$
\begin{aligned}
VN &= q\tau T(\pi E + \phi U) \text{ by } (17)\\
&= q\tau(\pi E + \phi U')T \text{ by } (20)\\
&= -(\pi' E + \phi' U')q\tau T \text{ by } (23)\\
&= -N'V.
\end{aligned}
$$

Type b. The characteristic equation of the matrix

(25) $$\pi = \begin{pmatrix} 0 & e \\ p & 0 \end{pmatrix}$$

is $p(\lambda^2) = 0$, so that π has the single elementary divisor $p(\lambda^2)$ and the matrix

(26) $$N = \pi E + U$$

has the single elementary factor $[p(\lambda^2)]^k$. If

(27) $$X = \begin{bmatrix} 0 & 0 & \cdot & 0 & -\epsilon \\ 0 & 0 & \cdot & \epsilon & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & (-\epsilon)^{k-1} & \cdot & 0 & 0 \\ (-\epsilon)^k & 0 & \cdot & 0 & 0 \end{bmatrix},$$

it is easily shown that

$$XU = -U'X$$

and that, if

(28) $$V = q\tau X,$$

V is non-singular and satisfies (18). It should be noted that, since τ is skew symmetric, V is symmetric, if k is even, and skew symmetric if k is odd.

Type c. If, in (26), $\pi = 0$ and $\epsilon = 1$, so that U is the auxiliary unit matrix of order k,

(29) $$N = U,$$

is a matrix with the single elementary factor λ^k. Moreover the matrix V defined by (28) where q, τ and ϵ all have the value unity, satisfies (18).

If $N_1, N_2, \cdots, N_r$ are r matrices, where N_i is the matrix N with $k = k_i$, the diagonal block matrix

(30) $$M = [N_1, N_2, \cdots, N_r]$$

has the elementary factors $[p(\lambda)]^{k_i}$, $[p(-\lambda)]^{k_i}$, if each N_i is defined by (17); the elementary factors $[p(\lambda^2)]^{k_i}$, if each N_i is defined by (26); and the ele-

mentary factors λ^{k_i}, if each N_i is defined by (29) ($i = 1$,
(24) or (28) with $k = k_i$ determines a non-singular
$V_i N_i = -N' V_i$ and accordingly the matrix

$$(31) \qquad Q = [V_1, V_2, \cdots, V_r]$$

is non-singular and satisfies (6), when M has the value

Let the elementary factors of $A - \lambda B$ be $[p_i$
($i = 1, 2, \cdots, s$), of type a, $[p_i(\lambda^2)]^{k_{ij}}$ ($i = s+1, \cdot$
$\lambda^{k_{tj}}$ of type c,

$$j = 1, 2, \cdots, r_i; \ k_{i1} \geqq k_{i2} \geqq \cdots \geqq$$

where $p_i(\lambda) \neq p_j(\lambda)$, $p_i(\lambda^2) \neq p_j(\lambda^2)$, if $i \neq j$. The

$$(32) \qquad M = [M_1, M_2, \cdots, M_t],$$

where M_i is the matrix corresponding to the matrix on
p is replaced by p_i and r by r_i, has the same elementar
the same invariant factors as $A - \lambda B$. Hence the mat
matrix AB^{-1}. Moreover, if Q_i is obtained from (31) in
obtained from (30), the diagonal block matrix,

$$(33) \qquad Q = [Q_1, Q_2, \cdots, Q_t],$$

satisfies the equation $QM = -M'Q$. We may accordi
M and Q defined by (32) and (33) as the matrices
Since any matrix G, commutative with M in (32) is a
$[G_1, G_2, \cdots, G_t]$ by the remark at the end of section
treat each block R_i separately.

3. *Reduction of R.* We consider the matrix $R_i =$
symmetric and non-singular, G_i is commutative with M_i
We first treat the case, in which $i \leqq t - 1$, so that t
of M_i are not of the form λ^k. For simplicity of writin
all suffixes i and write R, Q, G, M, etc. for R_i, Q_i, G_i,
M and Q are the matrices defined by equations (30)

If,

$$G = (G_{ij})$$

is a *partition* of G *similar* to that of M in (30), i. e. i

[6] R. C. Trott, *Bulletin of the American Mathematical*
Abstract No. 95.

the same number of rows k_i as N_i and the same number of columns k_j as N_j, it is known that, when $k_i \geqq k_j$,

$$G_{ij} = \begin{pmatrix} F_{ij} \\ 0 \end{pmatrix}, \qquad G_{ji} = (0 \;\; F_{ji}), \tag{34}$$

where F_{ij} and F_{ji} are square matrices of order $k = k_j$.[7] Moreover,

$$F_{ij} = f_{ij0}E + f_{ij1}U + \cdots + f_{ij,k-1}U^{k-1}, \tag{35}$$

where $f_{ijs} = f_{ijs}(\pi)$ is a polynomial in the matrix π with coefficients in K. Since $R = QG$ is skew-symmetric, it follows that

$$V_iG_{ij} = -G'_{ji}V'_j = \rho_jG'_{ji}V_j, \tag{36}$$

where $\rho_j = +1$, if V_j is skew symmetric and $\rho_j = -1$ if V_j is symmetric. In particular, if $k_i = k_j$, since $G_{ij} = F_{ij}$ and $V_i = V_j$, (36) becomes

$$V_iF_{ij} = \rho_jF'_{ji}V_j = \rho_iF'_{ji}V_i. \tag{37}$$

As a consequence of the definition of V_i,

$$V_i\pi U^a = -\pi' U'^a V_i.$$

Hence,

$$\begin{aligned} V_iF_{ij} &= V_i \sum_{s=0}^{k-1} f_{ijs}(\pi)U^s, \\ &= \sum_{s=0}^{k-1} f_{ijs}(-\pi')U'^sV_i, \\ &= \rho_i \sum_{s=0}^{k-1} f_{jis}(\pi')U'^sV_i \quad \text{by (37)}. \end{aligned}$$

Therefore, if $k_i = k_j$,

$$f_{ijs}(-\pi) = \rho_i f_{jis}(\pi) \qquad (s = 1, 2, \cdots, k-1). \tag{38}$$

In particular,

$$f_{iis}(-\pi) = \rho_i f_{iis}(\pi).$$

Hence $f_{iis}(\pi)$ is an even polynomial in π if $\rho_i = 1$ and an odd polynomial, if $\rho_i = -1$. In either case $f_{iis}(\pi)$ is singular, if and only if it is zero. Consequently we have the result; *G_{ii} is singular, if and only if its first element f_{ii0} is zero.*

Let $k_1 = k_2 = \cdots = k_c > k_{c+1} \geqq k_{c+j}$. Then, if G_{11} is singular, but for some value of j, $1 < j \leqq c$, G_{jj} is non-singular, we may interchange the first and j-th rows, and the first and j-th columns, thus bringing G_{jj} into the place

[7] Trott, *loc. cit.*, cf. Cullis, *Matrices and Determinoids*, vol. 3, chap. XXVII.

of G_{11} without disturbing M or Q. We therefore suppose that G_{jj} is singular for all values of j, $1 \leqq j \leqq c$. Since the first element of G_{i1} is zero, when $i > c$, (equation (34)) and G is non-singular, the first element f_{j10} of G_{j1} is different from zero for at least one value of j, $1 < j \leqq c$. Accordingly without any loss of generality we may suppose that $f_{210} \neq 0$.

Let

$$W = \left[\begin{pmatrix} E_1 & 0 \\ w(\pi)E_1 & E_1 \end{pmatrix}, E_2\right],$$

where E_1 is the unit matrix of the same order as N_1 and E_2 the unit matrix of the same order as $[N_3, N_4, \cdots, N_r]$. The matrix W is commutative with M and

$$W'Q = \left[\begin{pmatrix} E_1 & w(\pi')E_1 \\ 0 & E_1 \end{pmatrix}, E_2\right] Q,$$

$$= Q \left[\begin{pmatrix} E_1 & w(-\pi)E_1 \\ 0 & E_1 \end{pmatrix}, E_2\right].$$

If $W'QGW = QH$, where $H = (H_{ij})$ is a partition of H similar to that of G,

$$H_{11} = G_{11} + w(-\pi)G_{21} + w(\pi)G_{12} + w(-\pi)w(\pi)G_{22}.$$

The first element h of H_{11} accordingly satisfies the equation

$$h = f_{110} + w(-\pi)f_{210} + w(\pi)f_{120} + w(-\pi)w(\pi)f_{220}$$

or

$$h = w(-\pi)f_{210} + w(\pi)f_{120},$$

since by hypothesis $f_{110} = f_{220} = 0$.

If $w(\pi)$ is the identity matrix h has the value $f_{210} + f_{120}$ and if $w(\pi) = \pi$, the value $\pi(f_{120} - f_{210})$. Since π is non-singular both these values of h cannot be zero, as otherwise f_{210} would be zero, contrary to our assumption. Thus we find a non-singular matrix W, such that $W'QG = QH$ where the first element of H_{11} is non-zero so that H_{11} is non-singular. We may therefore suppose that such a transformation has already been applied to G and accordingly may assume that G_{11} is non-singular.

The matrices Q and G now satisfy the hypothesis of Lemma 1, so that G may be reduced to a form, in which $G_{j1} = G_{1j} = 0$, $j \neq 1$. By $r-1$ repetitions of the above process we finally reduce G to the diagonal block matrix,

$$G = [G_1, G_2, \cdots, G_r], \tag{39}$$

where $G_j = \sum_{s=0}^{k_j-1} \gamma_{js} U_j{}^s$ and $\gamma_{js} = \gamma_{js}(\pi)$ is a polynomial in the matrix π while γ_{j0} is non-singular.

We now proceed to reduce the matrix G_j to the form $\gamma_{j0}E_j$. Let

$$\gamma_{j1} = \gamma_{j2} = \cdots = \gamma_{jc} = 0, \quad \gamma_{j,c+1} \neq 0, \quad c \leqq k_{j-1}.$$

Then, if $W_j = E_j - \frac{1}{2}\gamma_{j0}^{-1}\gamma_{j\,c+1}U_j^{c+1}$,

$$W_j^2 G_j = H_j = \gamma_{j0}E_j + \sum_{s=c+2}^{k_j-1} h_{js}(\pi)U_j^s.$$

Moreover

$$\begin{aligned} W'_j V_j &= \{E_j - \tfrac{1}{2}(\gamma_{j0}^{-1}E_j)'(\gamma_{j\,c+1}U_j^{c+1})'\}V_j \\ &= \rho_j^2 V_j(E_j - \tfrac{1}{2}\gamma_{j0}^{-1}E_j\gamma_{j\,c+1}U_j^{c+1}), \text{ by (36)}, \\ &= V_j W_j. \end{aligned}$$

Therefore $$W'_j V_j G_j W_j = V_j W_j^2 G_j = V_j H_j.$$

But H_j is of the same type as G_j, except that it contains no term in U_j^{c+1}. Accordingly, in at most $k_j - 1$ such steps, we may reduce G_j to the form

$$G_j = \gamma_j E_j. \tag{40}$$

The matrix γ_j in (40) is a polynomial in π, which is even or odd according as V_j is skew symmetric or symmetric.

It is now necessary to distinguish between the two cases;

case a. $$\pi = \begin{pmatrix} p & 0 \\ 0 & -p \end{pmatrix},$$

case b. $$\pi = \begin{pmatrix} 0 & e \\ p & 0 \end{pmatrix}.$$

In case a for all values of j, V_j is skew symmetric, so that γ_j in (40) is an even polynomial

$$\gamma_j = g_j(\pi^2) = \begin{pmatrix} g_j(p^2) & 0 \\ 0 & g_j(p^2) \end{pmatrix}.$$

Let $$r(\pi) = \begin{pmatrix} [g_j(p^2)]^{-1} & 0 \\ 0 & e \end{pmatrix}.$$

Then

$$\begin{aligned} r(\pi)'q\tau g_j(\pi^2)r(\pi) &= q\tau\begin{pmatrix} e & 0 \\ 0 & [g_j(p^2)]^{-1} \end{pmatrix}\begin{pmatrix} g_j(p^2) & 0 \\ 0 & g_j(p^2) \end{pmatrix}\begin{pmatrix} [g_j(p^2)]^{-1} & 0 \\ 0 & e \end{pmatrix}, \\ &= q\tau\begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix} \text{ by (21) and (22).} \end{aligned}$$

Therefore, if $W_j = r(\pi)E_j$,

$$W'_j V_j G_j W_j = V_j \begin{pmatrix} e & 0 \\ 0 & e \end{pmatrix} E_j = V_j = q\tau T_j.$$

Hence in case a we may reduce R and RM by transformations permutable with M, to the forms

$$(41) \qquad [q_T T_1, q_T T_2, \cdots, q_T T_r], \qquad [q_T T_1 N_1, q_T T_2 N$$

respectively. The matrices (41) are uniquely determi
q and by the exponents k_i to which $p(\lambda)$ and $p(-$
mentary factors of $A - \lambda B$.

We condense the above results in the following s
to each pair of elementary factors $[p(\lambda)]^k$, $[p(-$
canonical form of, $A - \lambda B$, *there is a block*

$$(42) \qquad VN - \lambda V,$$

where N *is defined by* (17) *and* V *by* (24).

case b. In this case no such great simplificati
Let

$$G = [\gamma_1 E_1, \gamma_2 E_2, \cdots, \gamma_r E_r], \qquad H = [\sigma_1 E_1,$$

where σ_i and γ_i are all non-singular polynomials in
non-singular matrix commutative with M such tha

$$(43) \qquad W'QGW = QH.$$

If, as in previous cases, we write $W = (W_{ij})$ as a m
consequence of (43) that

$$(44) \qquad \sum_{a=1}^{r} W'_{ai} V_a \gamma_a E_a W_{aj} = \delta_{ij} V_i \sigma_i E_i$$

δ_{ij} the Kronecker δ.

Since W is commutative with M

$$W_{ij} = \begin{pmatrix} F_{ij} \\ 0 \end{pmatrix}; \quad W_{ji} = (0\ F_{ji}), \quad i \leqq j \quad \text{i}$$

where F_{ij} is defined by (35). Moreover,

$$\begin{aligned} F'_{ij} V_j &= \{f_{ij0}(\pi') E_j + \sum_{s=1}^{k_j - 1} f_{ijs}(\pi') U'^{s}_j\} \\ &= V_j \{f_{ij0}(-\pi) E_j + \sum_{s=1}^{k_j - 1} (-1)^s f \\ &= V_j \bar{F}_{ij}. \end{aligned}$$

Accordingly if $i \leqq j$,

$$W'_{ij} V_i = (F'_{ij}\ 0) V_i = (0\ F'_{ij} V_j) = (0\ V_j$$

and

$$W'_{ji}V_j = \begin{pmatrix} 0 \\ F'_{ji} \end{pmatrix} V_j = \begin{pmatrix} 0 \\ F'_{ji}V_j \end{pmatrix} = \begin{pmatrix} 0 \\ V_j\tilde{F}_{ji} \end{pmatrix} = V_i \begin{pmatrix} \tilde{F}_{ji} \\ 0 \end{pmatrix}.$$

Hence,

(45) $$W'_{ij}V_i = V_j\tilde{W}_{ij} \quad \text{and} \quad W'_{ji}V_j = V_i\tilde{W}_{ji}$$

where

$$\tilde{W}_{ij} = (0\ \tilde{F}_{ij}) \quad \text{and} \quad \tilde{W}_{ji} = \begin{pmatrix} \tilde{F}_{ji} \\ 0 \end{pmatrix} \quad k_i \geqq k_j.$$

Therefore, if $w_{ij}(\pi)$ and $\tilde{w}_{ij}(\pi)$ are the first elements of the matrices W_{ij} and $\tilde{W}_{ij}$ respectively we have the results

(46) $$\tilde{w}_{ij}(\pi) = w_{ij}(-\pi), k_i = k_j;\ \tilde{w}_{ij}(\pi) = 0, k_i > k_j;\ w_{ij}(\pi) = 0, k_i < k_j.$$

It follows from (44) and (45) that

$$V_i \sum_{a=1}^{r} \tilde{W}_{ai}\gamma_a E_a W_{aj} = \delta_{ij} V_i \sigma_i E_i,$$

or since V_i is non-singular that

(47) $$\sum_{a=1}^{r} \tilde{W}_{ai}\gamma_a E_a W_{aj} = \delta_{ij}\sigma_i E_i \qquad (i, j = 1, 2, \cdots, r).$$

As a consequence of the nature of the matrices $\tilde{W}_{ai}$ and W_{aj}, (47) remains true when each matrix is replaced by its first element, so that

$$\sum_{a=1}^{r} \tilde{w}_{ai}(\pi)\gamma_a w_{aj}(\pi) = \delta_{ij}\sigma_i \qquad (i, j = 1, 2, \cdots, r).$$

If $k_{c-1} > k_c = k_{c+1} = \cdots = k_d > k_{d+1}$, it follows from (46) and the last equation that

(48) $$\sum_{a=c}^{d} w_{ai}(-\pi)\gamma_a w_{aj}(\pi) = \sigma_i\delta_{ij}, \quad c \leqq i \leqq d, \quad c \leqq j \leqq d.$$

The matrices γ_i and σ_i $(i = c, c+1, \cdots, d)$, are either all even polynomials in π or else all odd polynomials in π. We may therefore write

(49) $$\gamma_i = g_i(\pi^2)\pi^a, \qquad \sigma_i = h_i(\pi^2)\pi^a \qquad (a = 0 \text{ or } 1),$$

so that (48) becomes,

(50) $$\sum_{a=c}^{d} w_{ai}(-\pi)g_a(\pi^2)w_{aj}(\pi) = h_i(\pi^2)\delta_{ij}, \quad c \leqq i \leqq d, \quad c \leqq j \leqq d.$$

If θ^2 is a zero of $p(x)$, the field $K_1 = K(\theta^2)$ is simply isomorphic to the

field of all polynomials in π^2 with coefficients in K and the field $K_2 = K(\theta)$ is simply isomorphic to the field of all polynomials in π with coefficients in K. Accordingly (50) implies

$$(51) \qquad \sum_{a=c}^{d} w_{ai}(-\theta) g_a(\theta^2) w_{aj}(\theta) = h_i(\theta^2)\delta_{ij},$$

and conversely (51) implies (50). The field K_2 is quadratic over K_1 and, if $w(\theta)$ is an element of K_2, $w(-\theta) = \bar{w}$, is its conjugate. Hence, if

$$C = (c_{ij}) \qquad (i, j = 1, 2, \cdots, d + 1 - c),$$

where

$$c_{ij} = w_{c+i-1, c+j-1}(\theta),$$

(51) is equivalent to

$$(52) \qquad C^*[g_c(\theta^2), g_{c+1}(\theta^2), \cdots, g_d(\theta^2)]\, C = [h_c(\theta^2), \cdots, h_d(\theta^2)],$$

where C^* is the conjugate transposed of C.

By a suitable interchange of rows and columns it can be shown that $| w_{ij}(\pi) |$, $c \leqq i, j \leqq d$ is a factor of $| W |$. Hence, since W is non-singular, $| w_{ij}(\pi) | \neq 0$ and therefore $| C | \neq 0$, so that C is non-singular. Hence, *the two matrices* $[g_c(\theta^2), g_{c+1}(\theta^2), \cdots, g_d(\theta^2)]$ *and* $[h_c(\theta^2), h_{c+1}(\theta^2), \cdots, h_d(\theta^2)]$ *with elements in* K_2 *are equivalent under a non-singular conjunctive transformation with elements in* K_1.

Conversely, if in (52) θ is replaced by π, we have

$$\tilde{C}(\pi)[g_c(\pi^2), \cdots, g_d(\pi^2)]C(\pi) = [h_c(\pi^2), \cdots, h_d(\pi^2)]$$

and, if W_c is the direct product of $C(\pi)$ and the unit matrix E_c,

$$W_c[g_c(\pi^2)E_c, \cdots, g_d(\pi^2)E_d]\, W_c = [h_c(\pi^2)E_c, \cdots, h_d(\pi^2)E_d].$$

Multiplying this last equation by π^a and using (49) we get,

$$(53) \qquad \tilde{W}_c[\gamma_c E_c, \cdots, \gamma_d E_d]\, W_c = [\sigma_c E_c, \cdots, \sigma_d E_d].$$

Equation (43) implies a set of equations (53), one for each distinct value of the exponents k_i of $p(\lambda^2)$. If $W = [W_1, W_2, \cdots, W_f]$ and $W_1, W_2, \cdots, W_f$ are the matrices W_c of (52) corresponding to the distinct equations of the set (53),

$$\tilde{W}GW = H \text{ or } Q\tilde{W}GW = QH. \quad \text{Since } Q\tilde{W} = W'Q,$$

it follows that $W'QGW = QH$ and, since each matrix W_c is non-singular,

that W is non-singular. Hence in case b we may reduce R and RM by transformations permutable with M to the forms,

$$(54) \qquad [q\tau\gamma_1 X_1, \cdots, q\tau\gamma_r X_r] \quad \text{and} \quad [q\tau\gamma_1 X_1 N_1, \cdots, q\tau\gamma_r X_r N_r]$$

respectively where $\gamma_i = g_i(\pi^2)$, if k_i is odd and $\gamma_i = \pi g_i(\pi^2)$, if k_i is even. The matrices (54) are not uniquely determined by the matrices p, q and the exponents k_i of $p(\lambda^2)$. We may express the results as follows: *If* $[p(\lambda^2)]^k$ *occurs exactly* a *times among the elementary factors of the pencil* $A - \lambda B$, *corresponding to* $[p(\lambda^2]^k$ *in the canonical form there is a block*

$$(55) \qquad [V\gamma_1 N - \lambda V\gamma_1, \cdots, V\gamma_a N - \lambda V\gamma_a]$$

where N *is defined by* (26) *and* V *by* (28). *With this block, and so with* $[p(\lambda^2)]^k$, *is associated a diagonal matrix of order* a *with elements in the field* $K(\theta^2)$, *where* θ^2 *is a zero of* $p(x) = 0$. *This associated matrix is not uniquely determined but is determined apart from a non-singular conjunctive transformation in the field* $K(\theta)$.

If K is the field of all real numbers the only irreducible even polynomials are of the type $p(\lambda^2) = \lambda^2 + b^2$. Hence $K(\theta^2) = K$ and $K(\theta)$ is the field of all complex numbers. Since any real quadratic form is equivalent in the real field to a sum of a certain number of positive and negative squares, the matrix associated with an elementary factor $(\lambda^2 + b^2)^k$ may be reduced to the simple form $[\rho_1, \rho_2, \cdots, \rho_a]$ where $\rho_i = +1 \;\; i \leqq d$, $\rho_i = -1 \;\; i > d$, and d is uniquely determined. In fact d is the index of the quadratic form.

Case c. The reduction in this case is similar in many respects to that of the previous cases. Equations (34) and (35) are true, where f_{ijs} is now an ordinary number and no longer a matrix. We assume that

$$k_1 = k_2 = \cdots = k_c > k_{c+j}.$$

If k_1 is even V_1 is skew symmetric and, by a proof exactly similar to that in case a or b, we can reduce the matrix (G_{ij}) $(i, j = 1, 2, \cdots, c)$, to a diagonal matrix $[g_1 E_1, g_2 E_2, \cdots, g_c E_c]$ where, as in case b, the diagonal matrix $[g_1, g_2, \cdots, g_c]$ is only determined to within a non-singular congruent transformation with elements in K. Therefore corresponding to an elementary factor λ^{2k} there is in the canonical form of $A - \lambda B$ a block

$$gX_{2k}U_{2k} - \lambda g X_{2k},$$

where $g \neq 0$ and K_{2k} is defined by (27) with $\epsilon = 1$, while U_{2k} is the auxiliary unit matrix of order $2k$. On rearranging the rows and columns of X_{2k} and

U_{2k} in the order $1, 3, \cdots, 2k-1, 2, 4, \cdots, 2k$ we find that X_{2k} and U_{2k} are equivalent respectively to

$$\begin{pmatrix} 0 & -q \\ q & 0 \end{pmatrix} = q_T \quad \text{and} \quad \pi = \begin{pmatrix} 0 & e \\ p & 0 \end{pmatrix}$$

where e is the unit matrix of order k and p is the auxiliary unit matrix of order k while q, which satisfies (21), is the matrix T of (19) when $\epsilon = 1$. Hence *an elementary factor* λ^{2k} *may be considered to be of type b where* $p(\lambda) = \lambda^{2k}$ *and* $p = U$.

If k_1 is odd, V_1 is symmetric and $f_{110} = 0$, so that G_{11} is singular. As in previous cases we may suppose that $f_{210} \neq 0$ and since $f_{120} = -f_{210}$, it is easily shown that $\begin{vmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{vmatrix} \neq 0$. By repeated applications of Lemma 1 it therefore follows that c must be even and that G may be reduced to the [8] diagonal block form

$$[H_1, H_2, \cdots, H_{c/2}],$$

where H_j is a square matrix of $2k_1$ rows of the same type as $H_1 = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$. It is not possible to reduce the matrix H_1, for example, to diagonal form. Accordingly we proceed as follows and consider the pencil

$$\begin{pmatrix} V_1 & 0 \\ 0 & V_1 \end{pmatrix} H_1 \left\{ \begin{pmatrix} U_1 & 0 \\ 0 & U_1 \end{pmatrix} - \lambda \begin{pmatrix} E_1 & 0 \\ 0 & E_1 \end{pmatrix} \right\}. \tag{56}$$

The elementary factors of this pencil are $\lambda^{k_1}, \lambda^{k_1}$. But with the notation of (15) the elementary factors of $\phi U_1 - \lambda\epsilon E_1$ are also $\lambda^{k_1}, \lambda^{k_1}$, if $e = 1$. Hence the pencil (56) is equivalent to

$$VG(N - \lambda E),$$

where $N = \phi U_1$ and V is defined by (24) with $e = 1$ and $q = 1$. The matrix G is permutable with N and, if G is considered as a matrix of two rowed matrices, every element to the left of the leading diagonal is zero so that,

$$G = G_0 + G_1 + G_2 + \cdots + G_{k-1},$$

where G_i is the matrix formed by the elements of G in the i-th diagonal to the right of the leading one. Moreover, since G is non-singular G_0 is non-singular. Further $G_i N = N G_i$ and $VG_i = -G'_i V' = G'_i V$. If $G_1 = G_2 = G_{f-1} = 0$ and $G_f \neq 0$, the matrix $W = E - 1/2 G_0^{-1} G_f$ is permutable with N and satisfies the equation

$$W'VGW = V(G_0 + H_{f+1} + \cdots H_{k-1})$$

[8] This is a well known result. See Turnbull and Aitken, *Canonical Matrices*, p. 137.

where H_j is of the same type as G_j. Hence we may assume that $G_1 = G_2 = G_{k-1} = 0$. Since G_0 is commutative with N and VG_0 is skew symmetric,

$$G_0 = g\,\epsilon\,E_1,$$

where g is an element of K. A reduction similar to that in case b shows that g may be taken as $+1$.

Hence, *if λ^k occurs among the elementary factors of $A - \lambda B$ and k is odd it must occur an even number, $2a$, of times. In the canonical form of $A - \lambda B$ occur a blocks of the nature*

$$\tau T(\phi U - \lambda\,\epsilon\,E).$$

It should be noted that the two elementary factors λ^k, λ^k are accordingly of type (a) where $p(\lambda) = \lambda$ and π is the zero matrix. We have accordingly proved the theorem

THEOREM I. *A canonical form for the pencil $A - \lambda B$, where A is symmetric and B is skew symmetric, under non-singular congruent transformation in K, is a diagonal block matrix, whose component blocks are given by equation* (42) *or equation* (55).

COROLLARY. *Necessary and sufficient conditions that two such pencils $A - \lambda B$ and $C - \lambda D$ be equivalent in K are that,*

(a) *the elementary factors of $A - \lambda B$ be the same as those of $C - \lambda D$.*

(b) *the matrix associated with each elementary factor of the type $[p(\lambda^2)]^k$ in a normal form of $A - \lambda B$ be equivalent under a conjunctive transformation to the corresponding matrix in a normal form of $C - \lambda D$.*

4. *Reduction of B.* Since B is non-singular and skew symmetric there exists a non-singular matrix P with elements in K such that

$$P'BP = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}, \tag{57}$$

where E is the unit matrix of order one half the order of B. We now proceed to find a canonical form for the pencil $A - \lambda B$, in which B is equivalent to the simple matrix on the right of (57). We start with the canonical form deduced in the previous sections and have in all to consider three cases.

Case a. Corresponding to the elementary factors $[p(\lambda)]^k$, $[p(-\lambda)]^k$ in the canonical form is the block $q\tau TN - \lambda q\tau T$ (equation (42)), where

$$\tau T = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdot & 0 & -e \\ 0 & 0 & 0 & 0 & \cdot & e & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & -e & \cdot & 0 & 0 \\ 0 & 0 & e & 0 & \cdot & 0 & 0 \\ 0 & -e & 0 & 0 & \cdot & 0 & 0 \\ e & 0 & 0 & 0 & \cdot & 0 & 0 \end{bmatrix}.$$

By rearranging the rows and columns of τT in the order $1, 3, 5, \cdots, 2k-1, 2, 4, \cdots, 2k$, we see that $\tau T \approx \begin{pmatrix} 0 & -T_e \\ T_e & 0 \end{pmatrix}$[9] where T_e is the matrix

$$\begin{bmatrix} 0 & 0 & \cdot & e \\ \cdot & \cdot & \cdot & \cdot \\ 0 & e & \cdot & 0 \\ e & 0 & \cdot & 0 \end{bmatrix}.$$

The same transformation applied to N shows that

$$N \approx \begin{pmatrix} L & 0 \\ 0 & -L \end{pmatrix},$$

where

$$(58) \qquad L = \begin{bmatrix} p & e & \cdot & 0 & 0 \\ 0 & p & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & p & e \\ 0 & 0 & \cdot & 0 & p \end{bmatrix}.$$

If $W = \begin{pmatrix} q^{-1}T_e & 0 \\ 0 & E \end{pmatrix}$,

$$W'q\begin{pmatrix} 0 & -T_e \\ T_e & 0 \end{pmatrix} W = \begin{pmatrix} q^{-1}T_e & 0 \\ 0 & E \end{pmatrix}\begin{pmatrix} 0 & -qT_e \\ qT_e & 0 \end{pmatrix}\begin{pmatrix} q^{-1}T_e & 0 \\ 0 & E \end{pmatrix} = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}.$$

Further

$$W'q\begin{pmatrix} 0 & -T_e \\ T_e & 0 \end{pmatrix}\begin{pmatrix} L & 0 \\ 0 & -L \end{pmatrix} W = \begin{pmatrix} 0 & L \\ qT_eLq^{-1}T_e & 0 \end{pmatrix} = \begin{pmatrix} 0 & L \\ L' & 0 \end{pmatrix},$$

since $qT_eL = L'qT_e$. Hence

$$q\tau TN - \lambda q\tau T \approx \begin{pmatrix} 0 & L \\ L' & 0 \end{pmatrix} - \lambda\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}.$$

It accordingly follows that, if $R_aM_a - \lambda R_a$ is the part of the canonical form of $A - \lambda B$ depending on elementary factors of type a, including those of type λ^k where k is odd,

$$(59) \qquad R_aM_a - \lambda R_a \approx \begin{pmatrix} 0 & F \\ F' & 0 \end{pmatrix} - \lambda\begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix},$$

[9] We use $\approx$ to denote "is equivalent to."

where

(60) $$F = [L_1, L_2, \cdots, L_w]$$

is a diagonal block matrix, which is the direct sum of all matrices L defined by (58), one for each pair of elementary factors $[p(\lambda)]^k$, $[p(-\lambda)]^k$ of type a.

case b. Corresponding to the elementary factor $[p(\lambda^2)]^k$ in the canonical form is the block $q\tau\gamma XN - \lambda q\tau\gamma X$ (equation 55). It is necessary to consider the cases in which k is even and in which k is odd separately. If $k = 2f$ is even, the matrix $V = q\tau X$ is skew symmetric, so that $\gamma = g(\pi^2)\pi$ is an odd polynomial in π. Hence

(61) $$q\tau\gamma = -\gamma' q\tau.$$

By rearranging the rows and columns in the order $1, 3, \cdots, 2f-1, 2, 4, \cdots, 2f$ we find that

$$X \approx \begin{pmatrix} 0 & -T \\ T & 0 \end{pmatrix} \quad \text{and} \quad N \approx \begin{pmatrix} \pi E & E \\ U & \pi E \end{pmatrix}$$

where T is defined by (19) and E and U by (17).

If $\psi = q\tau\gamma$,

$$\psi' = \gamma' q' \tau' = -\gamma' q\tau = q\tau\gamma \text{ by (61)},$$

so that ψ is symmetric. Accordingly, if $W = \begin{pmatrix} E & 0 \\ 0 & \psi^{-1}T \end{pmatrix}$,

$$W' q\tau\gamma XW \approx \begin{pmatrix} E & 0 \\ 0 & T\psi^{-1} \end{pmatrix} \begin{pmatrix} 0 & -\psi T \\ \psi T & 0 \end{pmatrix} \begin{pmatrix} E & 0 \\ 0 & \psi^{-1}T \end{pmatrix} = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix}.$$

Similarly,

$$W' q\tau\gamma XNW \approx \begin{pmatrix} E & 0 \\ 0 & \psi^{-1}T \end{pmatrix} \begin{pmatrix} 0 & -\psi T \\ \psi T & 0 \end{pmatrix} \begin{pmatrix} \pi E & E \\ U & \pi E \end{pmatrix} \begin{pmatrix} E & 0 \\ 0 & \psi^{-1}T \end{pmatrix},$$

$$\approx \begin{pmatrix} -\psi TU & -\psi T\pi\psi^{-1}T \\ \pi E & \psi^{-1}T \end{pmatrix}.$$

Hence

(62) $$q\tau\gamma XN \approx Z = \begin{pmatrix} -\psi TU & \pi' E \\ \pi E & \psi^{-1}T \end{pmatrix}.$$

For example if $k = 6$,

$$Z = \begin{bmatrix} 0 & 0 & 0 & \pi' & 0 & 0 \\ 0 & 0 & -\psi & 0 & \pi' & 0 \\ 0 & -\psi & 0 & 0 & 0 & \pi' \\ \pi & 0 & 0 & 0 & 0 & \psi^{-1} \\ 0 & \pi & 0 & 0 & \psi^{-1} & 0 \\ 0 & 0 & \pi & \psi^{-1} & 0 & 0 \end{bmatrix},$$

where $\pi = \begin{pmatrix} 0 & e \\ p & 0 \end{pmatrix}$ and $\psi = \begin{pmatrix} -qpg & 0 \\ 0 & qg \end{pmatrix}$ and $g = g(p)$ is a polynomial in p.

We may write the matrix Z of (62) in the form $Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z'_{12} & Z_{22} \end{pmatrix}$, where $Z_{11} = -\psi TU$ etc. Then if $R_{b_1}M_{b_1} - \lambda R_{b_1}$ is that part of the canonical form of $A - \lambda B$ depending on elementary factors of type b, i. e. on $[p(\lambda^2)]^k$, where k is even,

(63) $$R_{b_1}M_{b_1} - \lambda R_{b_1} \approx \begin{pmatrix} Y_{11} & Y_{12} \\ Y'_{12} & Y_{22} \end{pmatrix} - \lambda \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix},$$

where

(64) $$Y_{ij} = [Z_{ij,1}, Z_{ij,2}, \cdots, Z_{ij,w}] \qquad (i, j = 1, 2),$$

is a diagonal block matrix, which is the direct sum of all matrices $Z_{ij,r}$ one for each elementary factor $[p(\lambda^2)]^k$, k even.

If however k is odd, the matrix V in (55) is symmetric so that $\gamma = g(\pi^2)$ is an even polynomial in π. In fact $q\tau\gamma = \begin{pmatrix} 0 & -q \\ q & 0 \end{pmatrix}\begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix}$, where $g = g(p)$ is a polynomial in p and

$$q\tau\gamma X = \begin{pmatrix} 0 & 0 & \cdot & 0 & 0 & 0 & qg \\ 0 & 0 & \cdot & 0 & 0 & -qg & 0 \\ 0 & 0 & \cdot & 0 & -qg & 0 & 0 \\ 0 & 0 & \cdot & qg & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & qg & \cdot & 0 & 0 & 0 & 0 \\ -qg & 0 & \cdot & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Rearranging the rows and columns of this matrix in the order 1, 3, 5, etc., we find

$$q\tau\gamma X \approx \begin{pmatrix} 0 & -qgX_e \\ qgX_e & 0 \end{pmatrix},$$

where X_e is symmetric, and is defined by (27) with ϵ replaced by e. The same transformation shows that $N \approx \begin{pmatrix} U & E \\ pE & U \end{pmatrix}$. If $W = \begin{pmatrix} (qg)^{-1}X_e & 0 \\ 0 & E \end{pmatrix}$,

$$W'q\tau\gamma XW \approx \begin{pmatrix} X_e(qg)^{-1} & 0 \\ 0 & E \end{pmatrix}\begin{pmatrix} 0 & -qgX_e \\ qgX_e & 0 \end{pmatrix}\begin{pmatrix} X_e(qg)^{-1} & 0 \\ 0 & E \end{pmatrix} = \begin{pmatrix} 0 & -E \\ E & 0 \end{pmatrix},$$

while

(65) $$W'q\tau\gamma XNW \approx \begin{pmatrix} -p(qg)^{-1}X_e & -U \\ -U' & qgX_e \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = C.$$

For example, if $k = 3$,

$$C=\begin{pmatrix} 0 & 0 & p(qg)^{-1} & 0 & -e & 0 \\ 0 & -p(qg)^{-1} & 0 & 0 & 0 & -e \\ p(qg)^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -qg \\ -e & 0 & 0 & 0 & qg & 0 \\ 0 & -e & 0 & -qg & 0 & 0 \end{pmatrix}.$$

Accordingly, if $R_{b_2}M_{b_2}-\lambda R_{b_2}$ is that part of the canonical form of $A-\lambda B$ depending on elementary factors of type b, i. e. on $[p(\lambda^2)]^k$, where k is odd,

$$(66)\qquad R_{b_2}M_{b_2}-\lambda R_{b_2}\approx\begin{pmatrix} D_{11} & D_{12} \\ D_{12} & D_{22}\end{pmatrix}-\lambda\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix},$$

where

$$(67)\qquad D_{ij}=[C_{ij,1},C_{ij,2},\cdots,C_{ij,w}]\qquad (i,j=1,2),$$

is a diagonal block matrix which is the direct sum of all matrices C_{ij} one for each elementary factor $[p(\lambda^2)]^k$, k odd (including the case $\lambda^{2k}=p(\lambda^2)$). It is an immediate consequence of equations (59), (63) and (64) that

THEOREM 2. *The pencil $A-\lambda B$ is equivalent in K to the pencil*

$$A_1-\lambda\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix}$$

where

$$A_1=\begin{pmatrix} 0 & 0 & 0 & F & 0 & 0 \\ 0 & Y_{11} & 0 & 0 & Y_{12} & 0 \\ 0 & 0 & D_{11} & 0 & 0 & D_{12} \\ F' & 0 & 0 & 0 & 0 & 0 \\ 0 & Y'_{12} & 0 & 0 & Y_{22} & 0 \\ 0 & 0 & D'_{12} & 0 & 0 & D_{22}\end{pmatrix},$$

and F, Y_{ij}, D_{ij} are defined by (60), (64) and (67) respectively.

COROLLARY. *The symmetric matrix A is equivalent to the matrix A_1 under a non-singular congruent transformation in K, which leaves the skew symmetric matrix $\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix}$ invariant.*

This corollary follows immediately by substituting the matrix $\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix}$ for B in the pencil $A-\lambda B$.

5. *K the field of all real numbers.* In the canonical forms of the previous sections occur the matrices p and q, where p is any matrix with elements in K

whose characteristic equation is $p(\lambda)=0$, and q is a non-singular symmetric matrix satisfying the equation $qp=p'q$. If p is chosen as the companion matrix of $p(\lambda)=0$, a comparatively simple matrix q can be determined.[10] If however K is the field of all real numbers, there are only three possible types for the irreducible equation $p(\lambda)$, and the corresponding values of p and q are even more simple. These are

$$(1)\ p(\lambda)=\lambda-a;\quad (2)\ p(\lambda)=\lambda^2-2a\lambda+a^2+b^2;\quad (3)\ p(\lambda^2)=\lambda^2+a^2.$$

In case (1) $p=a$, $q=1$;

In case (2) $p=\begin{pmatrix} a & b \\ -b & a\end{pmatrix}$, $q=\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}$;

In case (3) $p=-a^2$, $q=1$, $\pi=\begin{pmatrix} 0 & 1 \\ -a^2 & 0\end{pmatrix}$.

Moreover each matrix g occurring in Y_{ij} or D_{ij} (equations (64) and (67)) now has the value ± 1.

The matrix $p=-a^2$, in case (3), is obtained by particularizing the general formula but for some purposes it is preferable to take $\pi=\begin{pmatrix} 0 & a \\ -a & 0\end{pmatrix}$ instead of $\begin{pmatrix} 0 & 1 \\ -a^2 & 0\end{pmatrix}$. If this is done, it is easily seen, that the matrix Z in (62) is unaltered, except that $\psi=\begin{pmatrix} ga & 0 \\ 0 & ga\end{pmatrix}$ where g is a real number. Since, in ψ, g may be replaced by any real number with the same sign, we may take $\psi=\begin{pmatrix} \rho & 0 \\ 0 & \rho\end{pmatrix}$ where $\rho=\pm 1$. Similarly the matrix C in (65) is only altered to the extent that

$$C_{11}=\rho a X_2;\qquad C_{22}=\rho a X_2.$$

In conclusion we exhibit the possible canonical forms, to one of which a real symmetric matrix A of order 4 can be reduced by a real non-singular congruent transformation, which leaves invariant the skew symmetric matrix $\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix}$. If A is non-singular, the possible elementary divisors of $A-\lambda\begin{pmatrix} 0 & -E \\ E & 0\end{pmatrix}$ are

(α) $(\lambda\pm a)$, $(\lambda\pm b)$; (β) $(\lambda\pm a\pm ib)$; (γ) $(\lambda\pm ia)$, $(\lambda\pm ib)$;
(δ) $(\lambda\pm a)$, $(\lambda\pm ib)$; (ϵ) $(\lambda\pm a)^2$; (ζ) $(\lambda\pm ia)^2$.

The corresponding canonical forms for A are

[10] J. Williamson, *loc. cit.*, p. 490.

$$(\alpha)\quad \begin{pmatrix} 0 & 0 & a & 0 \\ 0 & 0 & 0 & b \\ a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \end{pmatrix}; \qquad (\beta)\quad \begin{pmatrix} 0 & 0 & a & -b \\ 0 & 0 & b & a \\ a & b & 0 & 0 \\ -b & a & 0 & 0 \end{pmatrix};$$

$$(\gamma)\quad \begin{pmatrix} \rho a & 0 & 0 & 0 \\ 0 & \sigma b & 0 & 0 \\ 0 & 0 & \rho a & 0 \\ 0 & 0 & 0 & \sigma b \end{pmatrix} \rho, \sigma = \pm 1; \qquad (\delta)\quad \begin{pmatrix} 0 & 0 & a & 0 \\ 0 & \rho b & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho b \end{pmatrix}, \rho = \pm 1;$$

$$(\epsilon)\quad \begin{pmatrix} 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \\ a & 0 & 0 & 0 \\ 1 & a & 0 & 0 \end{pmatrix}; \qquad (\zeta)\quad \begin{pmatrix} 0 & 0 & 0 & a \\ 0 & 0 & -a & 0 \\ 0 & -a & \rho & 0 \\ a & 0 & 0 & \rho \end{pmatrix}, \rho = \pm 1.$$

The matrices in cases (α), (β), (ϵ) depend solely on the elementary divisors of the pencil; (γ) yields 4 or 3 non-equivalent matrices according as a is not or is the same as b while (δ) and (ζ) both yield two non-equivalent matrices.

If A is singular and the pencil has the pair of elementary divisors λ, λ the canonical form is obtained from (α) or (δ) by putting $a = 0$ and from (α) by putting $a = b = 0$ if the pencil contains the 4 elementary divisors $\lambda, \lambda, \lambda, \lambda$. If λ^2 occurs among the elementary divisors the canonical form is obtained from that corresponding to $(\lambda \pm ia)$ by replacing the first a by unity and the other by zero. If λ^4 is an elementary divisor the canonical form is

$$\begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\rho \\ 0 & 0 & -\rho & 0 \end{pmatrix}, \rho = \pm 1.$$

Thus we have determined a complete list of the possible canonical forms for the case $n = 4$.

THE JOHNS HOPKINS UNIVERSITY.

ON THE MOMENTUM PROBLEM FOR DISTRIBUTION FUNCTIONS IN MORE THAN ONE DIMENSION. II.

By E. K. Haviland.

It has recently been proved in a paper[1] which will be referred to as I and which is based on an extension of a method of M. Riesz that for the existence of a distribution function solving the momentum problem corresponding to a given n-dimensional matrix $\| c_{k_1 \ldots k_n} \|$ it is necessary and sufficient that the matrix be non-negative in the sense that if

$$P(x_1, \cdots, x_n) = \sum_{k_1=0}^{N_1} \cdots \sum_{k_n=0}^{N_n} a_{k_1 \ldots k_n} x_1^{k_1} \cdots x_n^{k_n}$$

be a polynomial non-negative for all $(x_1, \cdots, x_n)$, the corresponding functional value

$$P_c = \sum_{k_1=0}^{N_1} \cdots \sum_{k_n=0}^{N_n} a_{k_1 \ldots k_n} c_{k_1 \ldots k_n}$$

is likewise non-negative. A. Wintner has subsequently suggested that it should be possible to extend this result by requiring that the distribution function solving the problem have a spectrum contained in a preassigned set, a result which would show the well-known criteria for the various standard special momentum problems (Stieltjes, Herglotz, Hamburger, Hausdorff in one or more dimensions) to be but particular cases of the general n-dimensional momentum problem mentioned above. The purpose of the present note is to carry out this extension. It turns out that this unified and more general treatment of all these momentum problems is by no means more complicated than the several individual treatments to be found in the literature for these special cases. As in I, the proofs are given, for convenience, in the case of two dimensions.

The general result is given by the

Theorem. *For the existence of a distribution function $\phi(E)$ whose spectrum S is contained in a given set C of the plane and which is such that*

$$(1) \qquad \int\int_T x^n y^m d_{xy}\phi(E) = c_{nm}, \qquad (n, m = 0, 1, 2, \cdots ; c_{00} = 1),$$

[1] E. K. Haviland, "On the momentum problem for distribution functions in more than one dimension," *American Journal of Mathematics*, vol. 57 (1935), pp. 562-572. Further references to the momentum problem are to be found in this paper.

where T denotes the entire (x, y)-plane, it is necessary and sufficient that to any polynomial

$$P(x, y) = \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} x^n y^m,$$

non-negative for all points (x, y) of C, there correspond the non-negative functional value

$$P_c = \sum_{n=0}^{N} \sum_{m=0}^{M} a_{nm} c_{nm}.$$

The matrix $\| c_{nm} \|$ will then be said to be non-negative with respect to C.

Proof. We may suppose C (which may be the entire (x, y)-plane) closed, since a spectrum is necessarily closed and since polynomials non-negative on a set C are non-negative on its closure $\bar{C}$ also.

The necessity of the condition is immediately clear. For C can then be taken to be the spectrum S itself. As S is closed, it belongs to the domain of definition of ϕ. Then $\int\int_T = \int\int_{T-S} + \int\int_S$. The second integral vanishes and if $P(x, y) \geq 0$ on S, the last integral is non-negative. Hence if (1) is to hold, we must have $P_c \geq 0$.

The proof of the existence of a distribution function $\phi(E)$ satisfying (1) under the hypotheses of the present theorem is effectively identical with the proof of the sufficient condition of the theorem in the paper I. One has only to replace the expression "non-negative" by the expression "non-negative with respect to C." The functions $g_{ij}(x, y)$ are defined as before, but in extending the functional operation to the modul generated by finite linear combinations of $1, x, y, xy, \cdots; g_{11}(x, y), \cdots, g_{ij}(x, y)$, one has to consider those elements of preceding moduls not less than (or not greater than) $g_{ij}(x, y)$ for any (x, y) *in* C. Since functions non-negative in T, the (x, y)-plane, are a fortiori non-negative on C, the proof then follows without further changes and the existence of such a $\phi(E)$ is assured. It remains only to be shown that the spectrum of ϕ is contained in C.

Let $P: (\xi, \eta)$ be a point of $T - C$. As C is closed, $T - C$ is open. Hence it is possible to find among the everywhere dense set of lines $x = \xi_i$, $y = \eta_j$ (the notation being that of the paper I) four forming the sides of a rectangle $R_1: (\xi_a \leq x < \xi_b; \eta_c \leq y < \eta_d)$ which contains (ξ, η) in its interior and is in turn contained in the interior of $T - C$. Consider the functions

$$H_1(x, y) = g_{bc}(x, y) - g_{ac}(x, y) \quad \text{and} \quad H_2(x, y) = g_{bd}(x, y) - g_{ad}(x, y).$$

There exists a modul such that both functions belong to it (and hence to all

succeeding moduls). To the former function corresponds the functional value $\gamma_{bc} - \gamma_{ac}$ and to the latter the functional value $\gamma_{bd} - \gamma_{ad}$. Both are non-negative and they can differ only if $H_1(x, y) \neq H_2(x, y)$ for some point (x, y) of C. As the two are identical for all points of C (for they are identical everywhere outside R_1), we have $\psi(R_1) = \gamma_{bd} - \gamma_{ad} - \gamma_{bc} + \gamma_{ac} = 0$. Hence if R_2 is a non-singular rectangle of ψ lying in R_1 and containing (ξ, η), we see $\psi(R_2) = \phi(R_2) = 0$. Consequently,[2] (ξ, η), which was any point of $T - C$, does not belong to the spectrum of ϕ, i. e., the spectrum of ϕ is contained in C, q. e. d.

We shall now examine the criteria for the solubility of the various standard special momentum problems and shall show these to be particular cases of the general criterion contained in our theorem.

1. *The one-dimensional Hamburger problem.* It is known[3] that every polynomial $P(x, y)$ with real coefficients which is non-negative for all points on the x-axis can be written in the form

$$[A(x)]^2 + [B(x)]^2 + yF(x, y),$$

where $A(x)$, $B(x)$ and $F(x, y)$ are polynomials with real coefficients. To this will correspond by (1), the set C being now the real axis, a functional value of the form

$$\sum_{h=0}^{n} \sum_{k=0}^{n} a_h a_k d_{h+k} + \sum_{h=0}^{m} \sum_{k=0}^{m} b_h b_k d_{h+k},$$

since $c_{ij} = 0$, $j \neq 0$. Here, as in what follows, c_{n0} is denoted by d_n. This functional value will be non-negative if and only if every section of the matrix $\| d_{h+k} \|$, $(h, k = 0, 1, 2, \cdots)$, belongs to a non-negative definite quadratic form. This is the criterion for the solubility of the one-dimensional Hamburger momentum problem.

2. *The Stieltjes problem.* Again, since[4] every real polynomial $P(x, y)$ which is non-negative for all non-negative points on the x-axis is of the form

[2] Cf. E. K. Haviland, "On the theory of absolutely additive distribution functions," *American Journal of Mathematics*, vol. 56 (1934), p. 653.

[3] Cf., e. g., G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 2, p. 82. On the other hand, Hilbert has shown that it is not always possible to express a polynomial in n variables ($n \geqq 2$) as the sum of the squares of a finite number of polynomials, "Über die Darstellung definiter Formen als Summen von Formenquadraten," *Mathematische Annalen*, vol. 32 (1888), pp. 342-350; Cf. E. Artin, "Über die Zerlegung definiter Funktionen in Quadrate," *Hamburger Abhandlungen*, vol. 5 (1926), pp. 100-115.

[4] Cf. Pólya and Szegö, *ibid.*

$$[A(x)]^2 + [B(x)]^2 + x\{[C(x)]^2 + [D(x)]^2\} + yF(x, y),$$

where $A(x)$, $B(x)$, $C(x)$, $D(x)$ and $F(x, y)$ are polynomials with real coefficients, there will correspond to $P(x, y)$ by (1) a functional value of the form

$$\sum_{h=0}^{n} \sum_{k=0}^{n} a_h a_k d_{h+k} + \sum_{h=0}^{m} \sum_{h=0}^{m} b_h b_k d_{h+k} + \sum_{h=0}^{q} \sum_{k=0}^{q} f_h f_k d_{h+k+1} + \sum_{h=0}^{r} \sum_{k=0}^{r} g_h g_k d_{h+k+1},$$

since $c_{ij} = 0$, $j \neq 0$. This expression will be non-negative if and only if every section of each of the matrices $\| d_{h+k} \|$ and $\| d_{h+k+1} \|$, $(h, k = 0, 1, 2, \cdots)$, belongs to a non-negative definite quadratic form. This is the criterion for the solubility of the one-dimensional Stieltjes momentum problem.

3. *Case of the interval* $[-1, 1]$. In view of the Legendre polynomials, it, may, perhaps, be of interest to consider also the one-dimensional momentum problem associated with the interval $[-1, 1]$. Every real polynomial $P(x, y)$ which is non-negative on $(-1 \leq x \leq 1;\ y = 0)$ will be of the form [4]

$$[A(x)]^2 + (1 - x^2)\,[B(x)]^2 + yF(x, y),$$

where $A(x)$, $B(x)$ and $F(x, y)$ are polynomials with real coefficients. To this will correspond, since $c_{ij} = 0$, $j \neq 0$, a functional value of the form

$$\sum_{h=0}^{n} \sum_{k=0}^{n} a_h a_k d_{h+k} + \sum_{h=0}^{m} \sum_{k=0}^{m} b_h b_k (d_{h+k} - d_{h+k+2}).$$

This expression will be non-negative if and only if every section of each of the matrices $\| d_{h+k} \|$ and $\| d_{h+k} - d_{h+k+2} \|$, $(h, k = 0, 1, 2, \cdots)$, belongs to a non-negative definite quadratic form.

4. *The trigonometrical moment problem.* Again,[4] every real trigonometric polynomial $g(\vartheta)$ of degree n which is non-negative for all values of ϑ can be represented in the form $g(\vartheta) = |h(e^{i\vartheta})|^2$, where $h(z) = a_0 + a_1 z + \cdots + a_n z^n$. The existence of a distribution function $\phi(E) = \Phi(\vartheta)$ satisfying (1), which now takes the form

$$c_{nm} = \int_0^{2\pi} (\cos\vartheta)^n (\sin\vartheta)^m \, d\Phi(\vartheta),$$

together with the requirement that the functional value corresponding to $|h(e^{i\vartheta})|^2$ be non-negative, requires that in the matrix

$$\| \Gamma_{i-k} \|, \ (i, k = 0, 1, 2, \cdots), \quad \text{where} \quad \Gamma_n = \int_0^{2\pi} e^{in\vartheta} d\Phi(\vartheta),$$

every section belongs to a non-negative definite Hermitian form. This is,

in fact, the criterion for the solubility of the Herglotz trigonometric momentum problem.

5. *The one-dimensional Hausdorff momentum problem.* It is known [5] that any polynomial $f(x)$ non-negative in $[0,1]$ may be expressed as a linear combination with positive coëfficients of polynomials $x^m(1-x)^{p-m}$. Then any polynomial $P(x,y)$ non-negative in $(0 \leq x \leq 1;\ y=0)$ will be of the form $f(x) + yG(x,y)$, where $G(x,y)$ is a polynomial in x and y. Consequently, the functional value P_c corresponding to such a $P(x,y)$ will be non-negative if and only if the c_{ij} are such that $c_{ij}=0$, $j \neq 0$, and

$$\sum_{k=0}^{s} (-1)^k \binom{s}{k} d_{m+k} \geqq 0, \tag{2}$$

which is the criterion for the solubility of the one-dimensional Hausdorff momentum problem.

6. *The two-dimensional Hausdorff momentum problem.*[6] Any polynomial $P(x,y)$ non-negative in $(0 \leq x \leq 1;\ 0 \leq y \leq 1)$ may be expressed similarly in terms of polynomials $x^m y^n (1-x)^{p-m}(1-y)^{q-n}$, wherefore the functional value P_c corresponding to $P(x,y)$ will be non-negative if and only if the c_{ij} satisfy a condition analogous to (2). This condition is precisely the criterion for the solubility of the two-dimensional Hausdorff momentum problem.

7. *The two-dimensional Hamburger momentum problem.* The two-dimensional Hamburger momentum problem treated in I corresponds to the case where C is the entire (x,y)-plane and hence the theorem of that paper is a particular case of our present theorem.

THE JOHNS HOPKINS UNIVERSITY.

---

[5] Pólya and Szegö, *op. cit.*, vol. 2, p. 83, ex. 49. A polynomial non-negative in $[a,b]$ can be expressed as a linear combination with positive coefficients of polynomials $(x-a)^m(b-x)^{p-m}$. In the case referred to by Pólya and Szegö, $a=-1$ and $b=1$. In the present case, $a=0$ and $b=1$.

[6] Cf. T. H. Hildebrandt and I. J. Schoenberg, "On linear functional operations and the moment problem for a finite interval in one or several dimensions," *Annals of Mathematics*, ser. 2, vol. 34 (1933), pp. 317-328; also F. Hallenbach, *Zur Theorie der Limitierungsverfahren von Doppelfolgen*, Thesis (Bonn), 1933.

SOME REMARKS ON F. JOHN'S IDENTITY.

By Hans Rademacher.

Recently F. John [1] has proved the

Theorem. *If $f(x)$ is a periodic function of bounded variation with the period 1, and if $\gamma = p/q > 1$ is a given rational number, $(p, q) = 1$, then*

$$\sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n} f\left(x - \frac{\log n}{\log \gamma}\right) = \log \gamma \int_0^1 f(y)\, dy, \tag{1}$$

where $a_n(\gamma)$ is defined by

$$a_n(\gamma) = a_n(p/q) = \sum_{l=1}^{q} \exp\left[\frac{2\pi i n l}{q}\right] - \sum_{l=1}^{p} \exp\left[\frac{2\pi i n l}{p}\right]$$

or, which is the same,

$$a_n(\gamma) = \begin{cases} 0; & p \nmid n,\ q \nmid n, \\ -p; & p \mid n,\ q \nmid n, \\ q; & p \nmid n,\ q \mid n, \\ q - p; & p \mid n,\ q \mid n. \end{cases} \tag{2}$$

This interesting identity induces me to make the following three simple remarks, of which the first establishes a connection with the Riemann ζ-function, the second proves (1) for the wider realm of Riemann-integrable functions, the third gives a generalization of (1).

1. The most important special case of (1) is doubtless $f(x) = e^{2\pi i k x}$, k being an integer. If we put $\lambda^{-1} = \log \gamma$, we have to prove in this case

$$\sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n} \exp[2\pi i k (x - \lambda \log n)] = \lambda^{-1} \int_0^1 e^{2\pi i k y} dy$$

or

$$\sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n} \exp[-2\pi i k \lambda \log n] = \begin{cases} 0, & k \neq 0 \\ \lambda^{-1}, & k = 0. \end{cases} \tag{3}$$

We have

$$\sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n} \exp[-2\pi i k \lambda \log n] = \sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n^{1+2\pi i k \lambda}}. \tag{4}$$

But as by the definition (2) the sum $\sum_{n=1}^{N} a_n(\gamma)$ is bounded for all N, the series

$$Z(s) = \sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n^s} \tag{5}$$

[1] F. John, "Identitäten zwischen dem Integral einer willkürlichen Funktion und unendlichen Reihen," *Mathematische Annalen*, vol. 110, pp. 718-721.

is convergent for $\Re(s) > 0$ and defines there a regular analytic function of s. On the other hand it follows from (2) and (5) that for $\Re(s) > 1$

$$Z(s) = \sum_{q|n} \frac{q}{n^s} - \sum_{p|n} \frac{p}{n^s} = q^{1-s} \sum_{m=1}^{\infty} \frac{1}{m^s} - p^{1-s} \sum_{m=1}^{\infty} \frac{1}{m^s}$$

or

$$Z(s) = \zeta(s)\,(q^{1-s} - p^{1-s}). \tag{6}$$

The equation (6) holds for $\Re(s) > 0$. Now we have to distinguish two cases:

1) $k \neq 0$. We find from (6)

$$Z(1 + 2\pi ik\lambda) = \zeta(1 + 2\pi ik\lambda)\,(q^{-2\pi ik\lambda} - p^{-2\pi ik\lambda}).$$

But

$$q^{-2\pi ik\lambda} - p^{-2\pi ik\lambda} = \exp\left[-\frac{2\pi ik \log q}{\log p - \log q}\right] - \exp\left[-\frac{2\pi ik \log p}{\log p - \log q}\right] = 0,$$

since

$$\frac{\log q}{\log p - \log q} = \frac{\log p}{\log p - \log q} - 1. \tag{7}$$

Hence

$$Z(1 + 2\pi ik\lambda) = 0. \tag{8}$$

2) $k = 0$. In this case we have by (6)

$$Z(1) = \lim_{\epsilon \to 0} Z(1 + \epsilon) = \lim_{\epsilon \to 0} \zeta(1 + \epsilon)\,(q^{-\epsilon} - p^{-\epsilon}) = \lim_{\epsilon \to 0} \frac{q^{-\epsilon} - p^{-\epsilon}}{\epsilon},$$

$$Z(1) = -\log q + \log p = \log \gamma = \lambda^{-1}. \tag{9}$$

The formulae (4), (5), (8), (9) prove (3).

By means of a Fourier expansion, the equation (3) could, of course, be used to prove (1) for a rather extended class of functions $f(x)$. But this reasoning would involve some complications of convergence, which can be surmounted easily only for functions with absolutely convergent Fourier-series, e. g., functions with bounded derivative. However, instead of pursuing this method, we proceed to prove (1) directly in our next remark.

2. In order to study the expression

$$\lim_{N \to \infty} \sum_{n=1}^{\infty} \frac{a_n(\gamma)}{n} f(x - \lambda \log n)$$

for a Riemann-integrable function $f(y)$ it is obviously sufficient to consider only such N as are divisible by pq, since the $a_n(\gamma)$ and $f(y)$ are bounded. Now we have by (2)

$$\begin{aligned} S_M(f) &= \sum_{n=1}^{Mpq} \frac{a_n(\gamma)}{n} f(x-\lambda \log n) \\ &= \sum_{\substack{1 \leq n \leq Mpq \\ q \mid n}} \frac{q}{n} f(x-\lambda \log n) - \sum_{\substack{1 \leq n \leq Mpq \\ p \mid n}} \frac{p}{n} f(x-\lambda \log n) \\ &= \sum_{m=1}^{Mp} \frac{1}{m} f(x-\lambda \log mq) - \sum_{m=1}^{Mq} \frac{1}{m} f(x-\lambda \log mp) \\ &= \sum_{m=Mq+1}^{Mp} \frac{1}{m} f(x-\lambda \log m - \lambda \log q), \end{aligned}$$

upon making use of (7) and of the periodicity of $f(y)$. Moreover, there is no loss of generality in setting $x = \lambda \log q$, for if (1) is true for any special value x_0, it is also true for any other x, as $f(y)$ and $f(y-x_0+x)$, regarded as functions of y, are both periodic and Riemann-integrable. Hence all we have to prove is

$$\lim_{M\to\infty} S_M(f) = \lim_{M\to\infty} \sum_{m=Mq+1}^{Mp} \frac{1}{m} f(-\lambda \log m) = \lambda^{-1} \int_0^1 f(y)\,dy. \tag{10}$$

Now let us first treat the special "step-function"

$$\phi_\alpha(y) = \begin{cases} 1, & 0 \leqq y < \alpha \\ 0, & \alpha \leqq y < 1, \end{cases} \tag{11}$$

$\phi_\alpha(y)$ being defined in points outside the interval $0 \leqq y < 1$ by periodic repetition of (11) modulo 1. The parameter α is supposed to be such that $0 \leqq \alpha \leqq 1$; for the extreme values $\alpha = 0$ or 1 one of the two inequalities in (11) cannot be fulfilled. We have therefore $\phi_0(y) = 0$ and $\phi_1(y) = 1$ for all y.

For $f = \phi_\alpha$ (10) becomes

$$\lim_{M\to\infty} S_M(\phi_\alpha) = \lim_{M\to\infty} \sum_{m=Mq+1}^{Mp} \frac{1}{m} \phi_\alpha(-\lambda \log m) = \alpha\lambda^{-1}. \tag{12}$$

As this is trivially true for $\alpha = 0$, we can assume $0 < \alpha \leqq 1$. Now we have by (11)

$$\sum_{m=Mq+1}^{Mp} \frac{1}{m} \phi_\alpha(-\lambda \log m) = \sum_{\substack{Mq < m \leqq Mp \\ 0 \leqq -\lambda \log m < \alpha \pmod 1}} \frac{1}{m}, \tag{13}$$

where $0 \leqq x < \alpha \pmod 1$ means, of course, $0 \leqq x - [x] < \alpha$. But the conditions of summation on the right-hand side of (13) can be written

(a) $\qquad \lambda(\log M + \log q) < \lambda \log m \leqq \lambda(\log M + \log p),$

(b) $\qquad 1 - \alpha < \lambda \log m \leqq 1 \pmod 1.$

Condition (a) assigns to $\lambda \log m$ an interval of length $\lambda(\log p - \log q) = 1$. Thus of the infinite set of intervals (b) of length α, which are periodic

modulo 1, either just one falls in (a) or two parts of intervals of (b), together of length α, lie in (a), so that the summation in (13) is either of the type

$$\lambda(\log M + \log q) + u_M < \lambda \log m \leqq \lambda(\log M + \log q) + u_M + \alpha,$$

where $0 \leqq u_M \leqq 1 - \alpha$, or of the type

$$\begin{aligned}&\lambda(\log M + \log q) < \lambda \log m \leqq \lambda(\log M + \log q) + \beta_1\\ &\lambda(\log M + \log p) - \beta_2 < \lambda \log m \leqq \lambda(\log M + \log p), \qquad \beta_1 + \beta_2 = \alpha.\end{aligned}$$

Therefore (12) will be proved if we show that

$$(14) \qquad \lim_{M\to\infty} \sum_{\log M + v < \log m \leqq \log M + v + \beta\lambda^{-1}} \frac{1}{m} = \beta\lambda^{-1},$$

where $v = v_M$ may be any number lying between assigned bounds, $c \leqq v_M < C$. Now for $m > 1$

$$\int_m^{m+1} \frac{dt}{t} < \frac{1}{m} < \int_{m-1}^{m} \frac{dt}{t}$$

and consequently

$$\int_{Me^v}^{Me^v\gamma^\beta} \frac{dt}{t} < \sum_{Me^v < m \leqq Me^v\gamma^\beta} \frac{1}{m} < \frac{1}{Me^v} + \int_{Me^v}^{Me^v\gamma^\beta} \frac{dt}{t}$$

or

$$\beta < \sum_{Me^v < m \leqq Me^v\gamma^\beta} \frac{1}{m} < \frac{1}{Me^c} + \beta,$$

which proves (14) and therefore also (12).

Now any periodic step-function of period 1 can be built up as a linear combination of a finite number of step-functions $\phi_\alpha(y)$ of the special type (11) with different parameters α. Hence (10) is proved for arbitrary step-functions with a finite number of steps.

If, finally, $f(y)$ is a periodic Riemann-integrable function, we can, to any given $\epsilon > 0$, assign two step-functions $\phi(y)$ and $\Phi(y)$ of period 1, such that

$$(15) \qquad \phi(y) \leqq f(y) \leqq \Phi(y)$$

and

$$(16) \qquad \int_0^1 (\Phi(y) - \phi(y))\,dy < \epsilon.$$

Since (10) is valid for $\phi(y)$ and $\Phi(y)$, we have

$$(17) \qquad \begin{aligned}\lim_{M\to\infty} S_M(\phi) &= \lambda^{-1} \int_0^1 \phi(y)\,dy,\\ \lim_{M\to\infty} S_M(\Phi) &= \lambda^{-1} \int_0^1 \Phi(y)\,dy.\end{aligned}$$

But $S_M(f)$ shows in (10) only *positive* coefficients of $f(-\lambda \log m)$ and therefore we have from (15)

$$S_M(\phi) \leqq S_M(f) \leqq S_M(\Phi).$$

From this and (17) we conclude

$$\lambda^{-1} \int_0^1 \phi(y)\,dy \leqq \underline{\lim}_{M\to\infty} S_M(f) \leqq \overline{\lim}_{M\to\infty} S_M(f) \leqq \lambda^{-1} \int_0^1 \Phi(y)\,dy.$$

But according to (16) this proves (10) for any Riemann-integrable function $f(y)$, for which, therefore, John's identity (1) is true.

3. The relation between the identity (1) and the Riemann ζ-function, *discussed in* § 1, suggests the possibility of finding similar identities related to other ζ-functions.

Let K be a field of algebraic numbers, of degree n; let γ be a number of the field with $|N(\gamma)| > 1$ and

$$\gamma = \mathfrak{a}/\mathfrak{b}, \qquad (\mathfrak{a}, \mathfrak{b}) = 1.$$

Now for ideals $\mathfrak{n}$ of the field we introduce, in analogy with (2), the arithmetic function $a_{\mathfrak{n}}(\gamma)$ through the definition

$$(18) \qquad a_{\mathfrak{n}}(\gamma) = \begin{cases} 0 & \mathfrak{a} \nmid \mathfrak{n} \quad \mathfrak{b} \nmid \mathfrak{n} \\ -N(\mathfrak{a}) & \mathfrak{a} \mid \mathfrak{n} \quad \mathfrak{b} \nmid \mathfrak{n} \end{cases}$$

vergence of the series

$$Z(s) = \sum_{\mathfrak{n} \in \mathfrak{C}} \frac{a_{\mathfrak{n}}(\gamma)}{N(\mathfrak{n})^s}$$

can be proved only for $\Re(s) > 1 - [2/(n+1)]$, for which purpose we should have to use Landau's estimate [2] of the "ideal-function"

$$(20) \qquad H(x; \mathfrak{C}) = \sum_{\substack{\mathfrak{n} \in \mathfrak{C} \\ N(\mathfrak{n}) \leqq x}} 1 = \kappa \log x + O(x^{1-[2/(n+1)]}).$$

But instead of giving further details of this reasoning we prefer to pass immediately to the generalization of § 2, which is not quite so obvious.

For our proof we start with the remark that for a fixed A

$$\sum_{\substack{\mathfrak{n} \in \mathfrak{C} \\ MA < N(\mathfrak{n}) \leqq (M+1)A}} \frac{1}{N(\mathfrak{n})} \to 0 \quad \text{as} \quad M \to \infty.$$

Indeed, we have from (20)

$$\sum_{\substack{\mathfrak{n} \in \mathfrak{C} \\ MA < N(\mathfrak{n}) \leqq (M+1)A}} \frac{1}{N(\mathfrak{n})} \leqq \frac{1}{MA} \sum_{\substack{\mathfrak{n} \in \mathfrak{C} \\ MA < N(\mathfrak{n}) \leqq (M+1)A}} 1$$

$$= \frac{1}{MA} \kappa A + O\left(\frac{1}{M} M^{1-[2/(n+1)]}\right) = O(M^{-[2/(n+1)]}).$$

Hence for the study of

$$\mathfrak{C}' = \mathfrak{C}\mathfrak{C}_1^{-1}.$$

Hence we get, writing $\mathfrak{m}$ instead of $\mathfrak{m}_1$ and $\mathfrak{m}_2$,

$$S^*_M(f) = \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ N(\mathfrak{m})\leqq MN(\mathfrak{a})}} \frac{1}{N(\mathfrak{m})} f(x - \Lambda(\log N(\mathfrak{m}) + \log N(\mathfrak{b})))$$
$$- \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ N(\mathfrak{m})\leqq MN(\mathfrak{b})}} \frac{1}{N(\mathfrak{m})} f(x - \Lambda(\log N(\mathfrak{m}) + \log N(\mathfrak{a}))).$$

Because of $\Lambda^{-1} = \log |N(\gamma)| = \log N(\mathfrak{a}) - \log N(\mathfrak{b})$, we have

$$\Lambda(\log N(\mathfrak{m}) + \log N(\mathfrak{b})) - \Lambda(\log N(\mathfrak{m}) + \log N(\mathfrak{a})) = -1.$$

From this and the periodicity of $f(y)$ we conclude

$$S^*_M(f) = \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ MN(\mathfrak{b}) < N(\mathfrak{m})\leqq MN(\mathfrak{a})}} \frac{1}{N(\mathfrak{m})} f(x - \Lambda \log N(\mathfrak{m}) - \Lambda \log N(\mathfrak{b})).$$

As we saw in § 2, the choice of a special value for x involves no loss of generality. We put $x = \Lambda \log N(\mathfrak{b})$ and then have to prove

$$(21) \quad \lim_{M\to\infty} S^*(f) = \lim_{M\to\infty} \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ MN(\mathfrak{b}) < N(\mathfrak{m})\leqq MN(\mathfrak{a})}} \frac{1}{N(\mathfrak{m})} f(-\Lambda \log N(\mathfrak{m})) = \Lambda^{-1}\int_0^1 f(y)\,dy.$$

For the required proof we need the relation

$$(22) \qquad L(x) = \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ N(\mathfrak{m})\leqq x}} \frac{1}{N(\mathfrak{m})} = \kappa \log x + C + O(x^{-[2/(n+1)]}),$$

which follows from (20) by the customary process of Abel's partial summation. The constant C in (22) may depend on the class $\mathfrak{C}'$.

In complete analogy with § 2, we prove (21) only for the special step-function $\phi_a(y)$, defined in (11). We have

$$S^*_M(\phi_a) = \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ MN(\mathfrak{b}) < N(\mathfrak{m})\leqq MN(\mathfrak{a})}} \frac{1}{N(\mathfrak{m})} \phi_a(-\Lambda \log N(\mathfrak{m})) = \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ MN(\mathfrak{b}) < N(\mathfrak{m})\leqq MN(\mathfrak{a}) \\ 0\leqq -\Lambda\log N(\mathfrak{m}) < a \pmod 1}} \frac{1}{N(\mathfrak{m})}.$$

The conditions of this sum may be treated like those of (13), and our problem is then reduced to the proof of

$$\lim_{M\to\infty} \sum_{\substack{\mathfrak{m}\in\mathfrak{C}' \\ \log M+v < \log N(\mathfrak{m})\leqq \log M+v+\beta\Lambda^{-1}}} \frac{1}{N(\mathfrak{m})} = \kappa\beta\Lambda^{-1}.$$

According to (22) the left-hand side is equal to

$$\begin{aligned}
&\lim_{M\to\infty} \{L(Me^v \mid N(\gamma)\mid^\beta) - L(Me^v)\} \\
&= \lim_{M\to\infty} \{\kappa(\log M + v + \beta\Lambda^{-1}) - \kappa(\log M + v) + O(M^{-[2/(n+1)]})\} \\
&= \kappa\beta\Lambda^{-1},
\end{aligned}$$

which was to be proved.

The further arguments are quite the same as in § 2. We first consider arbitrary step-functions and can then enclose a given R-integrable function $f(y)$ between two step-functions $\phi(y)$ and $\Phi(y)$ whose integrals differ by as little as we wish. In this way the theorem of this paragraph is fully proved.

I close this article with a special example of the generalized John's identity (19). Let K be Gauss's field of complex numbers $a+bi$. We choose $\gamma = 1 + i$. As there is only the principal class of ideals, we can replace the ideals by integers $m + ni$ of the field. We have only to observe that each principal ideal is represented by four associated numbers. If we therefore sum over all integers (with the omission of 0), we get on both sides of (19) the four-fold amount. We notice further that $\kappa = \pi/4$ in this case and that

$$a_{\mathfrak{n}}(\gamma) = a_{m+ni}(1+i) = \begin{cases} 1, & m \not\equiv n \pmod 2 \\ -1, & m \equiv n \pmod 2, \end{cases}$$

or

$$a_{m+ni}(1+i) = (-1)^{m+n+1}.$$

Hence we have the equation

$$\sum_{m,n}{}' \frac{(-1)^{m+n+1}}{m^2+n^2} f\left(x - \frac{\log(m^2+n^2)}{\log 2}\right) = \pi \log 2 \int_0^1 f(y)\,dy,$$

the sum being extended over all pairs (m, n) with the omission of $(0, 0)$ and arranged according to increasing values of $m^2 + n^2$.

UNIVERSITY OF PENNSYLVANIA,
PHILADELPHIA, PA.

NOTE ON A THEOREM OF PONTRJAGIN.

By E. R. VAN KAMPEN.

I. The problem to determine under what conditions a locally euclidean group is a Lie group was solved for the compact case by v. Neumann [1] [1]. Later Pontrjagin [2] generalized his solution by proving the following theorem:

(1) *Any locally connected finite dimensional compact group is a Lie group.*[2]

In this note [3] we prove certain theorems related to (1) and implying (1). Some of these have already been stated in [5], section IX. The proofs follow easily from the structural considerations in [6] and similar more detailed results on abelian groups to be found in [3] and also in [4]. We wish to emphasize the central position occupied by Theorem 6 ([5], section IX, restated in [6]). This theorem, obtained as a corollary from the theory of almost periodic functions, can then be used as sole foundation of the structural theory of compact groups. While [5] has not been written with this idea in mind, an analysis of the line of thought used there will justify this statement for the case of abelian compact groups as well.

While we are considering compact separable groups, there is no difficulty at all in extending the results to bicompact groups.

II. *The closed invariant subgroups of a compact connected group.* By the same methods as used in [6], II-IV, we prove first:

(2) *Any closed invariant subgroup G of a compact connected group F can be generated by a closed subgroup of the centrum of F and certain of the invariant simple subgroups $S^{(l)}$ of F.*

Each compact Lie group F_n ([6], p. 301) contains a closed invariant subgroup G_n corresponding to G. For these groups (2) is well known. If for a certain number n and $l < p_n$. ([6], p. 302) the subgroup $S_n{}^{(l)}$ of F_n corresponding to $S^{(l)}$ is contained in G_n, then the same is true for all $m > n$, and $S^{(l)}$ is contained in G. Moreover G_n is generated by its centrum C'_n and

[1] The numbers in square brackets refer to the literature at the end of the note.

[2] In order to simplify the wording of certain statements, we allow Lie groups to have a finite number of components. In particular we include under that term as a degenerate case all finite groups.

[3] It can be considered as a continuation of [6] from which we take over all notations, in particular those of Theorems 1 and 2.

the subgroups $S_n^{(l)}$ contained in G_n. Here C'_n is the common part of G_n and the centrum C_n of F_n. Hence G is generated by its subgroups $S^{(l)}$ together with the common part of G and C.

III. *The dimension of a compact group.*

(3) *The dimension of a compact group F is equal to the sum of the dimensions of its simple Lie subgroups $S^{(l)}$ $(l = 1, 2, \cdots)$, and the dimension of its centrum C.*

We may of course suppose that F is connected and that C has the finite dimension m. If the number of groups $S^{(l)}$ is not finite, we can find locally euclidean sets of arbitrarily high dimension by considering the groups generated by certain finite collections of groups $S^{(l)}$. Thus we can assume that the number of groups $S^{(l)}$ is finite, and that they generate a Lie group S of dimension n. Considering that in this case the group A ([6], Theorem 1) is finite, so that F and F/A are locally homeomorphic, and applying the reasoning of [4], p. 458, h, on the centrum C, we find that a nucleus of F is homeomorphic with the product of an n-cell, an m-cell and a 0-dimensional set. Hence statement (3) is proved.

(4) *If H is a closed invariant subgroup of a compact group F, then the dimension of F is equal to the sum of the dimensions of H and F/H.*

We may obviously suppose that F is connected. Comparing (2) with (3) we see that (4) immediately reduces to the case of compact abelian groups. But in this case statement (4) follows from [4], p. 458, h.

(5) *Any sufficiently small subgroup H of a finite dimensional compact group F is 0-dimensional.*

We may again suppose that F is connected. Then F has a 0-dimensional invariant subgroup G with a Lie factorgroup. According to (3) and [6], Theorem 1, this is true for F, if it is true for the centrum C of F. But for C it is an immediate consequence of the properties of charactergroups (Compare [4]).

As a Lie group does not have arbitrarily small subgroups, we can restrict H to such an open set containing G that the image of H in F/G consists of the identity element only. But then H is contained in G, so that it is 0-dimensional.

The most noteworthy point about the proofs of (4) and (5) is that a thorough analysis of the structure of compact groups on the lines indicated

is at present unavoidable. As a result these simple sounding theorems cannot yet be proved for any class of groups for which no structural analysis is known. Important examples are the class of locally compact groups and the very restricted class of locally euclidean groups.

IV. In this section we consider a closed invariant subgroup G of a compact connected group F from the point of view of local connectedness. According to (2) the group G has subgroups S', C', A', related in the same way as the subgroups S, C, A of F respectively ([6], Theorem 1). The subgroup S' is generated by the simple Lie subgroups invariant in G, C' is the centrum of G and A' is the common part of S' and C'. According to [6], Theorem 3, the groups F and C/A are at the same time locally connected and not locally connected. Though G is not connected, Theorem 3 remains true for G and C'/A'. For C'/A' and G/S are simply isomorphic and S is connected. Hence there is a one-to-one correspondence between the components of G and the components of C'/A'.

By [6], Theorem 3, the components of the identity elements of G and C'/A' are locally connected and not locally connected at the same time. As we have just seen that G and C'/A' have at the same time a finite or an infinite number of components, it follows now that G and C'/A' are at the same time locally connected or not locally connected.

We consider the case that both F and C/A are locally connected, while G and C'/A' are not locally connected.

In the first place it is then possible that the factorgroup $A''' = A''/A'$ of A' in the common part A'' of C' and A is not locally connected. As A is 0-dimensional this means simply that A''' is not finite. Then A/A' is not finite, S/S' must be infinite dimensional and F/G must be infinite dimensional also.

In the second place we suppose that the factorgroup A''' is locally connected and accordingly finite. Then the factorgroup $C'' = C'/A''$ of the common part A'' of A and C' in C' is locally simply isomorphic with the factorgroup C'/A', so $C'' = C'/A''$. is not locally connected. The group C'' can be considered as a subgroup of the locally connected group C/A. Now a locally connected compact abelian group is the direct product of a certain collection of rotation groups. If such a group has a non-locally connected closed subgroup, the factorgroup cannot be finite dimensional. So we see:

(6) *If a closed invariant subgroup G of a locally connected compact group F is itself not locally connected then F/G is not finite dimensional.*

Pontrjagin's theorem (cited in (1)) is an immediate consequence of (5), (6) and the existence of arbitrarily small invariant subgroups with Lie factor-groups in any compact group. However, once (3) is proved (1) follows more easily by a direct argument:

In a finite dimensional compact group F the number of subgroups $S^{(l)}$ is finite, so that A is finite ([6], Theorem 1). But then, if F is also locally connected, C is locally connected ([6], Theorem 3) and as C is also finite dimensional it is a Lie group. Hence F is a Lie group.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] J. von Neumann, "Die Einführung analytischer Parameter in topologischen Gruppen," *Annals of Mathematics*, vol. 34 (1933), pp. 170-190.

[2] L. Pontrjagin, "Sur les groupes topologiques compacts," *Comptes Rendus*, vol. 198 (1934), p. 238.

[3] L. Pontrjagin, "The theory of topological commutative groups," *Annals of Mathematics*, vol. 35 (1934), pp. 341-388.

[4] E. R. van Kampen, "Locally bicompact abelian groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.

[5] E. R. van Kampen, "Almost periodic functions and compact groups," *Annals of Mathematics*, vol. 37 (1936).

[6] E. R. van Kampen, "The structure of a compact connected group," *American Journal of Mathematics*, vol. 57 (1935), pp. 301-308.

POINT SET THEORY APPLIED TO THE RANDOM SELECTION OF THE DIGITS OF AN ADMISSIBLE NUMBER.[1]

By Arthur H. Copeland.

Kamke and Tornier have pointed out objections to von Mises' definition of the collective (*Kollektiv*). The difficulty arises in connection with a selection operator (*Auswahl*) whose function is to transform one collective into another. In this paper we shall show how the difficulty can be overcome.

The collective is the fundamental element in von Mises' theory of probability.[2] It is the mathematical idealization of a sequence of physical occurrences (measurements, results of tossing a coin, etc.). A collective K consists of an infinite sequence of elements (points of some space S). Thus

$$K = e^{(1)}, e^{(2)}, e^{(3)}, \cdots \tag{1}$$

The points of S are called labels (*Merkmale*) and the space S is called the label space (*Merkmalraum*). Associated with every label is a probability defined as follows. Let m be any label belonging to S and let r_n be the number of times the label m occurs in the first n terms of K. Then r_n/n is called the success ratio for m in the first n trials of K and $\lim_{n\to\infty} r_n/n$ is called the probability of m with respect to K. The set of probabilities associated with the elements of S is called the distribution (*Teilung*). The first restriction imposed on the collective is that it must possess a distribution, i. e., the limit of the success ratio must exist for every element m of S.

The operation of "selection" is defined in the following manner. Let

$$n_1, n_2, n_3, \cdots \tag{2}$$

be any infinite increasing sequence of positive integers. We can form a new collective K' by selecting the n_1-st, n_2-nd, n_3-rd, $\cdots$ terms from $K = e^{(1)}, e^{(2)}, e^{(3)}, \cdots$. Thus $K' = e^{(n_1)}, e^{(n_2)}, e^{(n_3)}, \cdots$. It will be convenient to introduce a notation for this operation. Let us define the sequence

$$x = x^{(1)}, x^{(2)}, x^{(3)}, \cdots \tag{3}$$

[1] Presented to the Society April 11, 1930. This paper also contains the material presented Dec. 29, 1928, under the title *A proof that almost every number is admissible and is associated with the probability one-half.*

[2] See von Mises I and II. References to literature are given at the end of this paper.

such that the terms $x^{(n_1)}, x^{(n_2)}, x^{(n_3)}, \cdots$ are all 1's and the rest of the terms are all 0's. Then x is a sequence similar to K. The corresponding label space consists of the two points 1 and 0. It is also convenient to think of x as a number which lies between 0 and 1, is expressed in the binary scale, and has the digits $x^{(1)}, x^{(2)}, x^{(3)}, \cdots$. Any number x in the interval from 0 to 1 (0 excluded) can be a selection operator. In the case of the *ambiguous* representation, the ambiguity is decided by the condition that the sequence (2) is infinite and hence an infinite number of the digits of x must be equal to 1. The fact that K' is obtained by operating on K with x is expressed by the equation $K' = K \subset x$.[3]

The second restriction on the collective is that its distribution must be invariant under all selections which operate "*Ohne Benützung der Merkmalunterschiede*." Roughly this means that the distribution of a collective K is invariant under the operation of any selection which is random with respect to K. Von Mises regards as random any selection which is given by mathematical law independently of the collective. Kamke raises the following objection to this restriction. "*Es fragt sich, ob diese Genügsamkeit angebracht ist. Denn unterliegen die Folgen* (8)"[4] "*keinerlei Einschränkung, so kann man unabhängig von jeder E-Folge*"[4] (sequence (1)) "*die Gesamtheit aller folgen* (8) *bilden und für jede W-Folge*[4] *gäbe es dann unter diesen, unabhängig von der W-Folge gebildeten Index-folgen* (8), *stets eine solche, bei der auf jedes E_ν das Merkmal m zutrifft, wie auch eine solche, bei der auf kein E_ν das Merkmal m zutrifft. Dann gäbe es also überhaupt keine Kollektiv. Man darf also offenbar nicht beliebige Folgen* (8) *zulassen, sondern muß sich auf 'gesetzmäßige' oder 'mathematisch gegebene' Folgen* (8) *beschränken. Wie der Bereich dieser Folgen gegen die 'nicht gesetzmäßigen' Folgen abzugrenzen ist, diese Frage bleibt aber offen.*[5]

For the purpose of this paper it will be convenient to state Kamke's criticism in a slightly different form. To this end we shall replace the second

---

[3] The symbol $\subset$ is an inverted implication sign. If we think of K and x as both representing event sequences, then $K \subset x$ represents the event sequence "K if x," or "K is implied by x." (See Copeland III.) Dörge indicates the operation of selection by a product. (See Dörge I.) Since I have used the product $x \cdot y$ to indicate the event sequence x and y, it is necessary to use another symbol to indicate the event sequence "x if y." It is interesting to note that Dörge's *Einheit Auswahl* is represented in my notation by the number 1.

[4] Sequence (8) referred to by Kamke is the same as sequence (2) in my paper. An "E-Folge" is an arbitrary sequence K. An E-Folge is called a "W-Folge" with respect to a label m, provided, there exists a probability of m with respect to the sequence.

[5] Kamke I. Tornier points out a similar objection, Tornier I.

restriction on the collective by the following. The distribution of a collective must be invariant under the operation of every selection of a certain set E. We shall call the set E a *fundamental set* and for the present we shall leave it entirely undefined. Kamke points out that if the fundamental set consists of all selections which can be defined by mathematical law, then the set of collectives will be null.

Corresponding to any fundamental set, there arises the question of consistency of the two restrictions on the collective. We shall establish consistency by proving the existence of sequences satisfying the restrictions. It will be sufficient to consider a restricted type of sequence. We shall choose a label space which consists of the numbers 1 and 0. The sequences associated with this space admit of the same numerical representation as the selection operators. The reason for choosing such a restricted label space is that von Mises imposes certain other conditions with which this paper is not concerned but which are satisfied vacuously by the sequences associated with this label space. The space represents a simple alternative situation (heads or tails, etc.). We shall let the label 1 represent a success and the label 0 a failure. With this situation we can obtain the following simple expression for the success ratio.

$$p_n(K) = \sum_{i=1}^{n} (e^{(i)}/n) \tag{4}$$

where $p_n(K)$ is the success ratio of the label 1 in the first n trials of K. We shall denote the probability by $p(K)$ where $p(K) = \lim_{n\to\infty} p_n(K)$.

As an example of a fundamental set, we may take the set consisting of all selections whose numerical representations have the form

$$x_{r,n} = 2^{-r}/(1 - 2^{-n}) = 2^{-r} + 2^{-r-n} + 2^{-r-2n} \cdots,$$

where r and n are integers and $0 < r \leqq n$. Then the terms $x^r, x^{r+n}, x^{r+2n}, \cdots$ are all 1's, and the rest of the terms are all 0's. Hence, the operator $x_{r,n}$ selects the r-th, the $(r+n)$-th, the $(r+2n)$-th, $\cdots$ terms of K. It has been proved that there exists a set of sequences whose distributions are invariant under the operation of all selections of this fundamental set.[6] For these sequences, the associated label space consists of the elements 1 and 0. These sequences are called admissible numbers.[6] Admissible numbers also display the characteristics of the Bernoulli series, that is the probability of r successes in n trials is ${}_nC_r\, p^r(K)\,[1 - p(K)]^{n-r}$. In order that they may do

[6] For the proof of the existence of admissible numbers see Copeland I. A much simpler and more elegant proof has recently been given by von Mises. See von Mises III, example VI.

this, they must satisfy a certain independence condition which is defined in the following manner. Given two sequences K_1 and K_2, we can form a third sequence $K_1 \cdot K_2$ whose i-th term is a success (i. e., a 1) if and only if the i-th terms of both K_1 and K_2 are successes. Thus the i-th term of $K_1 \cdot K_2$ is the algebraic product of $e_1^{(i)}$ and $e_2^{(i)}$, and

$$K_1 \cdot K_2 = e_1^{(1)} \cdot e_2^{(1)}, e_1^{(2)} \cdot e_2^{(2)}, \; e_1^{(3)} \cdot e_2^{(3)}, \cdots \tag{5}$$

The sequences K_1 and K_2 are independent if and only if

$$p(K_1 \cdot K_2) = p(K_1) \cdot p(K_2).$$

The conjunction $K_1 \cdot K_2 \cdots K_n$ of the sequences $K_1, K_2, \cdots K_n$ is defined in the same manner as that of the conjunction of two sequences. A necessary and sufficient condition that n sequences $K_1, K_2, \cdots K_n$ be independent is that, for every subset $K_{r_1}, K_{r_2}, \ldots K_{r_\nu}$ of the set $K_1, K_2, \cdots K_n$,

$$p(K_{r_1} \cdot K_{r_2} \cdots K_{r_\nu}) = p(K_{r_1}) \cdot p(K_{r_2}) \cdots p(K_{r_\nu}).$$

A necessary and sufficient condition that a sequence K be an admissible number is that there exists a number $p (0 < p < 1)$ such that for every set of integers $r_1, r_2, \cdots r_\nu, n$ where $0 < r_1 < r_2 < \cdots < r_\nu \leqq n$,

$$p[(K \subset x_{r_1,n}) \cdot (K \subset x_{r_2,n}) \cdots (K \subset x_{r_\nu,n})] = p^\nu. \tag{6}$$

If K is an admissible number, then in particular $p(K \subset x_{r,n}) = p$. Furthermore, if r and n are both 1, then $x_{r,n}$ is the identity operator and $p(K) = p$. Thus K possesses a distribution and this distribution is invariant under all selections of the form $x_{r,n}$.[7] Moreover, the sequences

$$K \subset x_{1,n}, \; K \subset x_{2,n}, \cdots K \subset x_{n,n}$$

are independent.

Certain other properties of sequences can be invariant under the operation of selection. For example, it may happen that not only the distribution of a sequence is invariant under selection, but its property of being a collective is also invariant. Let $A(p)$ denote the set of all admissible numbers associated with the probability p. We shall say that *the properties of an admissible number K are invariant under the operation of a selection x, if $K \subset x$ belongs to the same set $A(p)$ as K.* I have proved that properties of all admissible numbers are invariant under all selections of the form $x_{r,n}$.

It will be observed that an admissible number is a collective whose fundamental set consists of the operators $x_{r,n}$. Thus, we already have one way of

[7] In my previous papers I have used the notation $(r/n)K$ instead of $K \subset x_{r,n}$.

getting around the difficulty raised by Kamke and Tornier. However, it might well be obejcted that this fundamental set is too restricted. For this reason we shall consider to what extent this set can be altered or augmented. At the conclusion of this paper we shall discuss other contributions to this problem.

We shall show that, corresponding to any denumerable fundamental set, there exists a continuum of admissible numbers whose properties are invariant under the operation of all selections of the fundamental set. A fundamental set can consist of almost every number from 0 to 1, "almost every" being used in the Lebesgue sense. The corresponding set of admissible numbers will then be at least denumerable. On the basis of the assumption of a well ordered continuum, we shall prove that both the fundamental set and the corresponding set of admissible numbers can be non-denumerable.

It seems to me sufficient to choose a denumerable fundamental set for the following reason. Let D be any denumerable fundamental set and let M be a set of admissible numbers. We shall show that M can contain almost every number in the interval from 0 to 1 and that if K is any member of M, then the properties of K will be invariant under the operation of every selection of D. Moreover, corresponding to any element K of M, there will exist a set E_K which contains almost every selection and all selections of which leave invariant the properties of K. It is important to notice that E_{K_1} and E_{K_2} are not necessarily identical if K_1 and K_2 are distinct.

The following theorem relates to the choice of the fundamental set.

THEOREM 1. *Given any denumerable set of selections D, there exists a set of admissible numbers M which has the power of the continuum and which is such that the operation of any selection of D on any admissible number of M, leaves invariant the properties of that admissible number.*

Let x be an arbitrary selection. Then $x = x^{(1)}, x^{(2)}, x^{(3)}, \cdots$ where $x^{(i)} = 1$ or 0. We shall let the set M consist of admissible numbers associated with the probability b/a, where a and b are integers such that $0 < b < a$. We shall let $y = y^{(1)}, y^{(2)}, y^{(3)}, \cdots$ where $y^{(i)} = 0, 1, 2, \cdots (a-1)$ and $K = e^{(1)}, e^{(2)}, e^{(3)}, \cdots$ where $e^{(i)} = 1$ if $y^{(i)} = 0, 1, \cdots (b-1)$, $e^{(i)} = 0$ otherwise.

We shall assume that a one to one correspondence has been established between the set of all positive integers λ and the set of all sets of integers $r_1, r_2, \cdots r_\mu, n$, such that $0 < r_1 < r_2 < \cdots < r_\mu \leqq n$. Then μ is defined as a function of λ. Let

$$U \equiv U(x, y, \lambda) = [(K \subset x) \subset x_{r_1,n}] \cdot [(K \subset x) \subset x_{r_2,n}] \cdot [(K \subset x) \subset x_{r_\mu,n}]$$

$$V \equiv V(x, y, \lambda) = \overline{\lim_{m\to\infty}} \, | \, p_m[U(x, y, \lambda)] - p_\lambda \, |$$

where $p_\lambda = (b/a)^\mu$ and $q_\lambda = 1 - p_\lambda$.

We shall prove that $V = 0$ almost everywhere in the region $\Delta: 0 < y < 1$. We have the equation

$$E(V \neq 0) = E(V > \tfrac{1}{2}) + E(V > \tfrac{1}{3}) + E(V > \tfrac{1}{4}) + \cdots.$$[8]

Hence, it will be sufficient to prove that $m[E(V > \epsilon)] = 0$ for every positive number ϵ. Since

$$E(V > \epsilon) = \lim_{m_0 \to \infty} \sum_{m=m_0}^{\infty} E[\,|\,p_m(U) - p_\lambda\,| > \epsilon],$$

it follows that

$$m[E(V > \epsilon)] \leqq \sum_{m=m_0}^{\infty} m\{E[\,|\,p_m(U) - p_\lambda\,| > \epsilon]\} \text{ for every } m_0.$$

This will imply

$$m[E(V > \epsilon)] \leqq \lim_{m_0 \to \infty} \sum_{m=m_0}^{\infty} m\{E[\,|\,p_m(U) - p_\lambda\,| > \epsilon\}.$$

We shall prove the convergence of this series. In order to do this, we have to compute the measure of the set $E[\,|\,p_m(U) - p_\lambda\,| > \epsilon]$. The expression $p_m(U)$ depends upon m digits of U and hence upon $m\mu$ digits selected from the first mn digits of $K \subset x$. The mn digits of $K \subset x$ are in turn selected from digits of K by means of the selection x. If ν is an integer such that $\nu \cdot p_\nu(x) = mn$, then the first mn digits of $K \subset x$ are selected from the first ν digits of K. Thus $p_m(U)$ is determined by the first ν digits of K and hence by the first ν digits of y.

The measure of the set of points y for which the first ν digits are prescribed, is $a^{-\nu}$. Our problem resolves itself into the counting of the number of permutations of the first ν digits of y which give rise to a U such that $|\,p_m(U) - p_\lambda\,| > \epsilon$.

Of the first ν digits of y, only mn are used in determining the first mn digits of $K \subset x$. Each of the remaining $\nu - mn$ digits has a possible values. Hence, there are $a^{(\nu - mn)}$ ways of selecting those digits which are not utilized. Of the first mn digits of $K \subset x$, there are only $m\mu$ digits which are used. The remaining $m(n - \mu)$ digits of K correspond to $m(n - \mu)$ digits of y which can be selected in $a^{m(n-\mu)}$ ways. Given a specified set of s digits of U which are equal to 1, and the remaining $n - s$ digits equal to 0, if $|\,p_m(U) - p_\lambda\,| > \epsilon$, then $|\,s/m - p_\lambda\,| > \epsilon$. The s digits of this specified set correspond to μs digits of y, each of which can be given b possible values. Hence, these digits of y can be selected in $b^{\mu s}$ ways. A digit of U which is 0, corresponds to μ digits of y which can be selected in $a^\mu - b^\mu$ ways. Thus there are $(a^\mu - b^\mu)^{m-s}$ ways in which the $\mu(m - s)$ digits of y can be chosen so that the specified set of $m - s$ digits of U will be 0. The s digits of U

[8] $E(V \neq 0)$ is the set of points for which $V \neq 0$.

which are equal to 1, can be selected in ${}_mC_s$ ways and s can take on all values consistent with the relations $| s/m - p_\lambda | > \epsilon$ and $0 \leqq s \leqq m$. Therefore

$$m\{E[p_m(U) - p_\lambda \mid > \epsilon]\} = \sum_{|s/m - p_\lambda| > \epsilon} {}_mC_s\, b^{\mu s}(a^\mu - b^\mu)^{m-s} a^{m(\nu-\mu)} a^{\nu - m\nu} a^{-\nu}$$
$$= \sum_{|s/m - p_\lambda| > \epsilon} {}_mC_s\, p_\lambda{}^s\, q_\lambda{}^{m-s}.$$

Since the series $\sum_{m=1}^{\infty} \sum_{|s/m-p_\lambda|>\epsilon} {}_mC_s\, p_\lambda{}^s\, q_\lambda{}^{m-s}$ converges,[9] it follows that

$m[E(V > \epsilon)] = 0$, and hence $V = 0$ almost everywhere in Δ.

Let the set D consist of the selections $x_1, x_2, \cdots$ and let x_0 be the identity selection. Thus $x_0 = 1 = 1, 1, 1, 1, \cdots$. We shall let the set M consist of the numbers K which correspond to the numbers y belonging to the set

$$C\{\sum_{i=0}^{\infty} \sum_{\lambda=1}^{\infty} E[V(x_i, y, \lambda) \neq 0]\} \qquad (C \text{ meaning “complement”})$$

Then every selection belonging to D leaves invariant the properties of every admissible number belonging to M. Since the measure of the set

$$\sum_{i=0}^{\infty} \sum_{\lambda=1}^{\infty} E[V(x_i, y, \lambda) \neq 0] \text{ is } 0,$$

the measure of its complement is 1. The set of numbers y associated with a given number K is of measure 0, and hence, the set M cannot be denumerable. If in particular, $a = 2$ and $b = 1$, then $K = 1 - y$ and the set M has the power of the continuum.

Theorem 1 shows that if the fundamental set be increased in such a way that it remains denumerable, this increase will not alter the set of admissible numbers appreciably. It also shows that in general the properties of an admissible number are not altered by the operation of selection. This fact will be brought out from another point of view by theorem 2.

The resultant of the operations of two selections on a collective is equivalent to the operation of a single selection on that collective. A fundamental set should be so chosen that the resultant of any two selections of the set, is itself a selection of the set. The fundamental set for admissible numbers possesses this group property.[10] It should be observed that if a fundamental set is increased, the set of collectives is not necessarily decreased. For example, the fundamental set for admissible numbers can be increased so as to include all rational selections.[11]

---

[9] For the proof of the convergence of this series, see Borel I, Chapitre I (Nombres normaux).

[10] A selection does not in general possess a unique inverse, and hence these transformations do not form a group.

[11] See Copeland I, theorem 17.

In theorem 2, x, y, K, $U(x,y,\lambda)$, and $V(x,y,\lambda)$ will be defined in the same manner as in theorem 1.

THEOREM 2. *For almost every y, $K \subset x$ is a member of the set $A(b/a)$ for almost every x (where $A(b/a)$ is the set of admissible numbers associated with the probability b/a).*

We shall prove first that $V(x,y,\lambda) = 0$ almost everywhere in the region Δ: $0 < x < 1$, $0 < y < 1$. We have to prove the convergence of the series

$$\sum_{m=m_0}^{\infty} m\{E[\,|\,p_m(U) - p_\lambda\,| > \epsilon]\}.$$

We have the relation

$$E[\,|\,p_m(U) - p_\lambda\,| > \epsilon] < E[p_{3mn}(x) \geqq 1/3] \cdot E[p_m(U) - p_\lambda\,| > \epsilon] \\ + E[p_{3mn}(x) < 1/3].$$

If $p_{3mn}(x) \geqq 1/3$, then there exists an integer, ν, such that $\nu p_\nu(x) = mn$ and $\nu \leqq 3mn$. Since $m\{E[p_{3mn}(x) \geqq 1/3]\} \leqq 1$, we have the inequality

$$m\{E[\,|\,p_m(U) - p_\lambda\,| > \epsilon]\} \leqq \sum_{|s/m - p_\lambda| > \epsilon} {}_mC_s\, p_\lambda{}^s\, q_\lambda{}^{m-s} + \sum_{\nu < mn} {}_{3mn}C_\nu \cdot 2^{-3mn}.$$

Hence, the series converges and $V = 0$ almost everywhere in Δ.

Let $E = \sum_{\lambda=1}^{\infty} E[V(x,y,\lambda) \neq 0]$. Then $m(E) = 0$ and E can be included in a set E' of Borel measure 0. Let $\phi(x,y)$ be the characteristic function of the set E'.[12] Then[13]

$$0 = \iint_\Delta \phi(x,y)\,dx\,dy = \int_0^1 dx \int_0^1 \phi(x,y)\,dy.$$

Therefore, $\int_0^1 \phi(x,y)\,dx = 0$ for almost every y, and if y is such that $\int_0^1 \phi(x,y)\,dx = 0$, then $\phi(x,y) = 0$ for almost every x. Thus for almost every y, $K \subset x$ is a member of the set $A(b/a)$ for almost every x.

It is easily seen that almost every point of Δ is such that K is a member of the set $A(b/a)$, and hence,

THEOREM 3. *For almost every y, $K \subset x$ is a member of the set $A(b/a)$ and K is a member of the set $A(b/a)$ for almost every x.*

The following theorem is a corollary of theorem 3.

[12] See de la Vallée Poussin I.
[13] See de la Vallée Poussin II.

THEOREM 4. *There exists a set of selections E and a set of admissible numbers M, such that E has the measure 1 and M is at least denumerable and the properties of every admissible number of M are invariant under every selection of E.*

Next, we shall show that both the set of selections E and the corresponding set of admissible numbers M can be non-denumerable. Let us assume that the numbers in the interval from 0 to 1 can be well ordered. Let us consider two such well ordered series and let one of them be called the selection series and the other, the admissible number series. The admissible number series shall contain all of the admissible numbers and no numbers which are not admissible. Consider the first member of the selection series. If the set of admissible numbers whose properties are not invariant under the operation of this selection, is not of measure 0, then this selection will be deleted from the series. Otherwise, we shall delete from the admissible number series all members whose properties are not invariant under the operation of the first member of the selection series. In either case, the remaining series will still be well ordered. Next let us consider the first element in the new admissible number series. This element will de deleted, if the set of selections which do not leave the properties of this element invariant is not of measure 0. Otherwise we shall delete from the selection series those selections which do not leave the properties invariant. This process will be continued, alternating between the selection series and the admissible number series. The order of procedure is determined, except for those elements which have no immediate predecessors. In the case of these elements, we shall perform the elimination for the selection series first. This process can not terminate in a denumerable number of steps. Hence we have

THEOREM 5. *There exists a set of selections E and a set of admissible numbers M, such that E and M are both nondenumerable, and the properties of every admissible number of M are invariant under the operation of every selection of E.*

It will be recalled that if K is admissible, then the numbers $K \subset x_{1,n}$, $K \subset x_{2,n}, \cdots K \subset x_{n,n}$ are independent. We shall consider to what extent this independence can be generalized. We shall, however, restrict ourselves to selections which are mutually exclusive.[14] This question can be investigated by means of the following device. Let x be expressed in the scale of n, i. e.,

[14] Two selections (considered as sequences) are mutually exclusive if their product is the sequence $0, 0, 0, \cdots$

$$x = x^{(1)}, x^{(2)}, x^{(3)}, \cdots \text{ where } x^{(k)} = 0, 1, 2, \cdots n-1 \qquad (k = 1, 2, \cdots).$$

Let

$$v_i = v_i^{(1)}, v_i^{(2)}, v_i^{(3)}, \cdots \text{ where } v_i^{(k)} = \begin{cases} 1 \text{ if } x^{(k)} = i-1, \\ 0 \text{ otherwise.} \end{cases}$$

Then the numbers $v_1, v_2, \cdots v_n$ represent n mutually exclusive selections. Given an admissible number K, we shall investigate the independence of the numbers $K \subset v_1, K \subset v_2, \cdots K \subset v_n$. We shall define the admissible number K by means of the equations

$$y = y^{(1)}, y^{(2)}, y^{(3)}, \cdots, \text{ where } y^{(i)} = 0, 1, 2, \cdots a-1,$$

and

$$K = e^{(1)}, e^{(2)}, e^{(3)}, \cdots, \text{ where } e^{(i)} = \begin{cases} 1 \text{ if } y^{(i)} = 0, 1, \cdots b-1, \\ 0 \text{ otherwise.} \end{cases}$$

Let

$$U(x, y, \lambda) = (K \subset v_{r_1}) \cdot (K \subset v_{r_2}) \cdots (K \subset v_{r_\mu})$$
$$V(x, y, \lambda) = \overline{\lim_{m \to \infty}} \mid p_m[U(x, y, \lambda) - p_\lambda \mid.$$

The scale n, in which x is expressed, depends upon λ, but x itself will be considered independent of λ. If $V(x, y, \lambda) = 0$, then the corresponding number K is said to satisfy a *generalized condition of admissibility*. If the set $\sum_{\lambda=1}^{\infty} E[V(x, y, \lambda) \neq 0]$ is of 0 measure, then K is said to satisfy almost every generalized condition of admissibility. We shall prove the following theorem.

THEOREM 6. *There exists a nondenumerable set of numbers such that each number K of the set satisfies almost every generalized condition of admissibility.*

We have the relation

$$E[\mid p_m(U) - p_\lambda \mid > \epsilon] < E[\mid p_m(U) - p_\lambda \mid > \epsilon] \cdot \prod_{i=1}^{\mu} E[p_{2mn}(v_{r_i}) \geqq 1/2n] + \sum_{i=1}^{\mu} E[p_{2mn}(v_{r_i}) < 1/2n].$$

If $p_{2mn}(v_{r_i}) \geqq 1/2n$, then there exists a ν_i such that $\nu_i \cdot p_{\nu_i}(v_{r_i}) = m$ and $\nu_i \leqq 2mn$. Let ν be the largest of the integers $\nu_1, \nu_2, \cdots \nu_\mu$.
Then

$$m\{E[\mid p_m(U) - p_\lambda \mid > \epsilon]\} \leqq \sum_{|s/m - p_\lambda| > \epsilon} {}_mC_s\, p_\lambda^s\, q_\lambda^{m-s} + \sum_{s/2mn < 1/2n} {}_{2mn}C_s\, (1/n)^s\, (1 - 1/n)^{2mn-s}.$$

The remainder of the proof is similar to that of theorem 3.

The theorems which we have proved indicate the possible latitude of choice for the fundamental set of selections. We shall now discuss briefly Tornier's contribution to the problem raised by Kamke.

Tornier replaces the linear sequences of von Mises by square matrices of the form $(e_i^{(j)})$ where the indices i and j take on separately all positive integral values. In place of the selection operation for the collective, Tornier permits the selection of rows from his matrices. He demands the independence of these rows where independence has the connotation previously mentioned in this paper. Tornier proves the consistency of such matrices. The disadvantage in his theory lies in the fact that it is not as accurate a picture of a set of physical measurements. In a set of measurements, linear order is indicated by time. A square array is formed only by mathematical artifice. Tornier's matrices are important in the following type of interpretation. Each row of such a matrix can be regarded as the sequence of measurements by a given experimenter. However, for such an interpretation, we should expect that each row of a matrix should possess properties similar to those of the collective. Such a matrix can be constructed. I have proved the existence of a set I of independent admissible numbers such that for every $p(0 < p < 1)$, $I \cdot A(p)$ has the power of the continuum.[15] The set of sequences I constitutes a matrix with a denumerable number of columns and a continuum of rows. Any ordered denumerable subset of I constitutes a Tornier matrix. The matrix I possesses the further property that its rows are collectives whose fundamental set consists of the selections $x_{r,n}$. The label space for I is, of course, the restricted space for admissible numbers. In a recent paper I have constructed a Tornier matrix such that the label space for each row has the power of the continuum, the probability distributions being given by Stieltjes integrals.[16]

REFERENCES.

E. Borel.

I. *Traité du Calcul des Probabilités*, T. II, Fasc. I, Ch. 1 (Paris, 1926).

A. H. Copeland.

I. "Admissible numbers in the theory of probability," *American Journal of Mathematics*, vol. 50, No. 4 (1928).

II. "Independent event histories," *American Journal of Mathematics*, vol. 51, No. 4 (1929).

III. "The theory of probability from the point of view of admissible numbers," *The Annals of Mathematical Statistics*, Aug. (1932).

IV. "A matrix theory of measurement," *Mathematische Zeitschrift*, Band 37 (1933).

[15] See Copeland II.

[16] See Copeland IV.

K. Dörge.

I. "Zu der R. v. Mises gegebenen Begründung der Wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, Band 32 (1930).

E. Kamke.

I. "Über neuere Begründungen der Wahrscheinlichkeitsrechnung," *Jahresbericht der Deutschen Mathematiker-vereinigung*, Band 42, Heft 1/4 (1932).

II. *Einführung in die Wahrscheinlichkeitstheorie.*

C. de la Vallée Poussin.

I. "Sur l'intégrale de Lebesgue," *Transactions of the American Mathematical Society* (1915).

II. *Intégrales de Lebesgue.*

R. von Mises.

I. "Grundlagen der Wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, Band 5 (1919).

II. *Vorlesungen aus dem Gebiete der angewandten Mathematik.*

III. "Zahlenfolgen mit kollektiv-ähnlichem Verhalten," *Mathematische Annalen*, Band 108 (1933).

E. Tornier.

I. "Wahrscheinlichkeitsrechnung und Zahlentheorie," *Journal für reine und angewandte Mathematik*, Band 160 (1929).

II. "Die Axiome der Wahrscheinlichkeitsrechnung," *Journal für reine und angewandte Mathematik*, Band 163 (1930).

The University of Michigan.

DEFINITION OF POST'S GENERALIZED NEGATIVE AND MAXIMUM IN TERMS OF ONE BINARY OPERATION.

By DONALD L. WEBB.

In 1921 Post[1] demonstrated that it was possible to construct a function for any order table in a system of m truth-values by the use of two primitive functions, $\sim_m p$ and $p \vee_m q$ which are generalizations of the functions $\sim p$ and $p \vee q$ in the two-valued case. Recently we[2] have been able to show that a function on m truth-values for any order table can be constructed in terms of one binary operation, using in this demonstration a negative that corresponds to Post's $\sim_m p$, a binary operator $p \, \alpha_m \, q$ which, for the value combinations used in the interpolation formula, corresponds to Post's $p \vee_m q$, and a binary operator $p \mid q$ which has no equivalent among the operators employed by Post. In the latter paper all operators were defined in terms of $p \mid q$. In this paper by redefining the truth-table of $p \mid q$ we are enabled to define Post's $\sim_m p$ and $p \vee_m q$ in terms of the " $\mid$ " function, thus greatly simplifying the proof that any m-valued logic can be generated by one binary operation. We find too that $p \mid q$ as so defined reduces in the two-valued case to one of Sheffer's functions,[3] as it evidently must.

The notation used in this paper is patterned after that of Post so as to avoid confusion.

Let $t_0, t_1, \cdots, t_{m-1}$, where m is any positive integer, signify the m truth values that an elementary proposition can assume in a m-valued logic. Denote by p, q elementary propositions. Let $p = t_i$ signify that the proposition p has the truth-value t_i. Make the two additional arithmetical definitions:

$$\begin{aligned} \min(i, j) &= i \quad \text{if} \quad i \leqq j \qquad\qquad (i, j = 0, 1, 2, \cdots) \\ &= j \quad \text{if} \quad i \geqq j; \\ i \equiv i_n \bmod n, &\quad (i = 0, 1, 2, \cdots) \qquad 0 \leqq i_n < n. \end{aligned}$$

Hence, $p \mid q$ is defined: if $p = t_i$, $q = t_j$ $(i, j = 0, 1, \cdots, m-1)$, then $p \mid q = t_k$ where $k = [\min(i, j) + 1]_m$.

[1] E. L. Post, *American Journal of Mathematics*, vol. 43 (1921), pp. 163-185.

[2] D. L. Webb, *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 252-254.

[3] H. M. Sheffer, *Transactions of the American Mathematical Society*, vol. 14 (1913), pp. 481-488.

THEOREM 1. $\sim_m p \equiv p \mid p$.

If $p = t_i$, then $p \mid p = t_k$ where $k = (i + 1)_m$. Thus $p \mid p$ cyclically permutes the truth-values t_i, giving $p \mid p$ and $\sim_m p$ the same truth-table. Therefore the two are equivalent.

Using Post's definition, $\sim_m^2 p = \sim_m \sim_m p$, etc., we may write

THEOREM 2. $p \vee_m q \equiv \sim_m^{m-1} (p \mid q)$.

By repeating the above process we find that if $p = t_i$, $\sim_m^h p = t_k$, where $k = (i + h)_m$ $(h = 2, 3, \cdots, m - 1)$. Hence, if $p = t_i$, $q = t_j$, then $\sim_m^{m-1} (p \mid q) = t_k$ where $k = \{[\min(i, j) + 1]_m + m - 1\}_m$, or $k = \min(i, j)$. But $p \vee_m q$ [4] as given by Post has the same truth-table, making the two equivalent.

Since Post has shown that we can generate a function of any order in a m-valued truth system by means of $\sim_m p$ and $p \vee_m q$, then, by using the above theorems, we can generate a function of any order in a m-valued truth system in terms of "$\mid$".

CALIFORNIA INSTITUTE OF TECHNOLOGY.

[4] This is called a maximum since the higher truth-value has the smaller subscript.

AN OPERATIONAL SOLUTION OF THE MAXWELL FIELD EQUATIONS.[1]

By E. P. Northrop.

1. *Introduction.* In empty space, containing no charges or currents, the classical electrodynamic field equations can, by a proper choice of units, be expressed as follows:

$$
\begin{aligned}
&\nabla \cdot \boldsymbol{e} = 0, \qquad && \nabla \cdot \boldsymbol{h} = 0 \\
&\nabla \times \boldsymbol{e} = -\,\partial \boldsymbol{h}/\partial t, \qquad && \nabla \times \boldsymbol{h} = \partial \boldsymbol{e}/\partial t.
\end{aligned} \tag{1.1}
$$

Here $\boldsymbol{e}$ is the electric intensity, and $\boldsymbol{h}$ the magnetic intensity. In the present paper we are concerned for the most part with the two equations involving the curls of $\boldsymbol{e}$ and $\boldsymbol{h}$. The significance of the other two equations will be discussed toward the end of the paper.

It is convenient for our purposes to think of the electromagnetic field as characterized by the six-component vector $\boldsymbol{v} = (e_x, e_y, e_z, h_x, h_y, h_z)$. This enables us to write the two curl equations of (1.1) as the single matrix equation

$$
\partial \boldsymbol{v}/\partial t = i\boldsymbol{H}\boldsymbol{v}, \tag{1.2}
$$

where $\boldsymbol{H}$ is the matrix operator

$$
\boldsymbol{H} = \left\{\begin{array}{c|c}
0 & \begin{matrix} 0 & -\frac{1}{i}\frac{\partial}{\partial z} & \frac{1}{i}\frac{\partial}{\partial y} \\ \frac{1}{i}\frac{\partial}{\partial z} & 0 & -\frac{1}{i}\frac{\partial}{\partial x} \\ -\frac{1}{i}\frac{\partial}{\partial y} & \frac{1}{i}\frac{\partial}{\partial x} & 0 \end{matrix} \\
\hline
\begin{matrix}\text{upper right,} \\ \text{signs changed}\end{matrix} & 0
\end{array}\right\}, \tag{1.3}
$$

and $\boldsymbol{v}$ is written as a one-column matrix. The solution of (1.2) is formally

$$
\boldsymbol{v} = e^{it\boldsymbol{H}}\boldsymbol{v}_0, \tag{1.4}
$$

[1] Presented to the Society September 6, 1934.

where $\boldsymbol{v}_0$ is the vector giving the initial state of the field. That is, the vector characterizing the electromagnetic field at any point and at any time can be expressed as an operator applied to $\boldsymbol{v}_0$. The purpose of this paper is to obtain rigorously an explicit integral form for the solution (1.4) of the equation (1.2). The treatment of the problem, and the terminology to be used will be based upon definitions and methods devised by M. H. Stone and others; and constant reference will be made to Stone's treatise "Linear Transformations in Hilbert Space," (*American Mathematical Society Colloquium Publications*, vol. 15, 1932). We shall refer throughout to this as simply "Stone."[2]

The reader, if unacquainted with the terminology of this work, would do well to refer to it for the definitions of the following terms: linear manifold (Definition 1.3), linear manifold determined by a set (1.4), transformation (2.1), extension of a transformation (2.2), adjoint of a transformation (2.8), symmetric transformation (2.9), self-adjoint transformation (2.11), essentially self-adjoint transformation (2.12), and unitary transformation (2.18). The following theorems are also of basic importance: Theorems 1.24, 1.25, 2.2, 2.6, 2.16, and 3.10. In addition, the discussion on unitary invariance at the end of the second chapter is worthy of attention.

It is perhaps advisable to make a few remarks in connection with these references. The space of functions which we shall use is the space $L_{2,6}$, composed of all vector point-functions $\boldsymbol{f}$ with components $(f_1, \cdots, f_6)$ defined over the whole of Euclidean space of three dimensions, and belonging to L_2.[3] The operations $+$ and $\cdot$ are defined as vector addition and scalar multiplication, the null element is defined to be $(0, \cdots, 0)$; and the function $(\boldsymbol{f}, \boldsymbol{g})$ is determined by the equation

$$(\boldsymbol{f}, \boldsymbol{g}) = \int\!\!\int_{-\infty}^{+\infty}\!\!\int (f_1\bar{g}_1 + \cdots + f_6\bar{g}_6)\, dxdydz,$$

where the bar denotes complex conjugate. This space is a special case of the space which is shown to be a Hilbert space in Theorem 1.25. The norm of a

---

[2] The author's attention has been called to a series of three articles by G. Herglotz in the *Berichte der Sächsischen Akademie*, vols. 78 (1926) and 80 (1928). In these articles (see in particular section 11, part III), methods are devised which, if modified and applied to the two curl equations of (1.1), appear to lead to results similar to those obtained in the present paper. These methods, however, neglect completely the questions of convergence and of domains of applicability of the various operators employed.

[3] It should be clearly understood that the functions to be considered in this article are functions of the time, t, as well as of the space coördinates x, y, z. The variable t, however, will be suppressed; and we shall write simply $f(x, y, z)$, etc.

function is denoted by $|f|$, and is defined as $(f,f)^{1/2}$. Throughout the paper we shall mean $\int_{-A}^{A}\int_{-A}^{A}\int_{-A}^{A}$ when we write $\int\int\limits_{-A}^{A}\int$.

The Fourier transformation is introduced in Theorem 3.10, but in a rather general way. The form which we shall have occasion to use can be described briefly as follows. Let T denote the Fourier transformation, and let $f(x,y,z)\,\varepsilon\, L_2$. Then

$$Tf(x,y,z) = \underset{A\to\infty}{\text{l.i.m.}}\, \frac{1}{(2\pi)^{3/2}} \int\int\limits_{-A}^{A}\int e^{i(x\xi+y\eta+z\zeta)} f(\xi,\eta,\zeta)\, d\xi d\eta d\zeta.$$

The symbol l.i.m. signifies as usual the limit in the mean (here of order 2). That is, $f(x,y,z) = \underset{n\to\infty}{\text{l.i.m.}}\, f_n(x,y,z)$ if $|f - f_n| \to 0$ as $n\to\infty$. Recall that if $f\,\varepsilon\, L_2$, then Tf and $T^{-1}f$ also belong to L_2. $T^{-1}f$ is defined by the same expression as above, save that the kernel is replaced by its complex conjugate.

2. *Outline of procedure.* Our problem, as stated in the last section, is to determine the operator $\boldsymbol{F}(\boldsymbol{H}) \equiv e^{it\boldsymbol{H}}$. A brief outline of the procedure to be followed may be of aid in understanding what will later be taken up in detail.

If $\boldsymbol{H}$ is a self-adjoint transformation, then $\boldsymbol{F}(\boldsymbol{H})$ is determined [4] by means of the equation

$$(\boldsymbol{F}(\boldsymbol{H})f, g) = \int_{-\infty}^{+\infty} F(\lambda)\, d(\boldsymbol{E}(\lambda)f, g).$$

The transformation $\boldsymbol{E}(\lambda)$ appearing in this relation is obtained by means of a certain contour integral [5] involving the inverse $\boldsymbol{H}_l^{-1}$ of the transformation $\boldsymbol{H}_l \equiv \boldsymbol{H} - l\boldsymbol{I}$, where $\boldsymbol{I}$ is the identity transformation, and l is an arbitrary not-real number. Thus the normal procedure would seem to be as follows: given the self-adjoint transformation $\boldsymbol{H}$, calculate $\boldsymbol{H}_l$, $\boldsymbol{H}_l^{-1}$, $\boldsymbol{E}(\lambda)$ and finally, $\boldsymbol{F}(\boldsymbol{H})$. Due, however, to the difficulties involved in the manipulation of the partial differential operator $\boldsymbol{H}$, we shall follow a less direct route.

The transform of the operator $\boldsymbol{H}$ by the Fourier transformation—call it $\boldsymbol{T}$—leads to a relatively simple algebraic operator, which we shall denote by $\boldsymbol{H}'$. That is $\boldsymbol{H}' \equiv \boldsymbol{THT}^{-1}$; and by Stone, Theorem 2.55, $\boldsymbol{H}'$ is self-adjoint. We then calculate in succession $\boldsymbol{H}'_l$, $\boldsymbol{H}'^{-1}_l$, and $\boldsymbol{E}'(\lambda)$. But $\boldsymbol{H}' \equiv \boldsymbol{THT}^{-1}$ implies [6] $\boldsymbol{E}'(\lambda) \equiv \boldsymbol{TE}(\lambda)\boldsymbol{T}^{-1}$ which is equivalent to $\boldsymbol{E}(\lambda) \equiv \boldsymbol{T}^{-1}\boldsymbol{E}'(\lambda)\boldsymbol{T}$. Hence we are enabled to calculate $\boldsymbol{E}(\lambda)$, and so, $\boldsymbol{F}(\boldsymbol{H})$.

[4] Stone, Theorem 6.1.
[5] Stone, Theorem 5.10.
[6] Stone, Theorem 7.1.

3. *The operators* $\boldsymbol{H}$ *and* $\boldsymbol{H}'$. The matrix operator $\boldsymbol{H}$ has already been defined in (1.3). The unitary transformation by which we obtain the transform of $\boldsymbol{H}$ is the diagonal matrix $\boldsymbol{T}$ whose elements are T, the Fourier transformation. The inverse $\boldsymbol{T}^{-1}$ of $\boldsymbol{T}$ is obtained by replacing T by T^{-1} in $\boldsymbol{T}$. It is then easy to show formally that the transform of $\boldsymbol{H}$ by $\boldsymbol{T}$ is the operator

$$(3.1)\qquad \boldsymbol{H}' \equiv \boldsymbol{T}\boldsymbol{H}\boldsymbol{T}^{-1} \equiv \left\{ \begin{array}{c|c} \mathbf{0} & \begin{matrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{matrix} \\ \hline \begin{matrix}\text{upper right,} \\ \text{signs changed}\end{matrix} & \mathbf{0} \end{array} \right\}.$$

It is evident that this operator lends itself much more readily to manipulation than does $\boldsymbol{H}$. It is to be noted, however, that we have only formally defined $\boldsymbol{H}$ and $\boldsymbol{H}'$ since we have said nothing of their domains.

It is relatively easy to find a domain in which $\boldsymbol{H}'$ is self-adjoint. On the other hand, the problem of showing that $\boldsymbol{H}$ is the inverse transform of $\boldsymbol{H}'$ (i. e., $\boldsymbol{H} \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{T}$, in which case the self-adjointness of $\boldsymbol{H}$ is established, and its domain determined) presents complications. Let us indicate a possible method of solving this problem in the form of a theorem.

THEOREM I.

Hypothesis. 1.) *Let* $\boldsymbol{H}'$ *be a self-adjoint transformation with domain* $D(\boldsymbol{H}')$. 2.) *Denote* $\boldsymbol{H}'$ *restricted to a domain* $D(\boldsymbol{H}'_0) \subset D(\boldsymbol{H}')$ *by* $\boldsymbol{H}'_0$; *and let* $\boldsymbol{H}'_0$ *be essentially self-adjoint.* 3.) *Define* $\boldsymbol{H}_0$ *by means of the relations* $\boldsymbol{H}_0 \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'_0\boldsymbol{T}$, $D(\boldsymbol{H}_0) \equiv \boldsymbol{T}^{-1}D(\boldsymbol{H}'_0)$. 4.) *Denote the adjoint of* $\boldsymbol{H}_0$ *by* $\boldsymbol{H}^*_0$ *and let* $\boldsymbol{H}^*_0 \equiv \boldsymbol{H}$, *where the domain of* $\boldsymbol{H}$ *is* $D(\boldsymbol{H})$.

Conclusion. $\boldsymbol{H}$ *is self-adjoint, and* $\boldsymbol{H} \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{T}$ *throughout*

$$D(\boldsymbol{H}) \equiv \boldsymbol{T}^{-1}D(\boldsymbol{H}').$$

Proof. Since $\boldsymbol{H}'_0$ is essentially self-adjoint, so also is $\boldsymbol{H}_0$. Hence $\boldsymbol{H}^*_0 \equiv \boldsymbol{H}$ is self-adjoint. But by Stone, Theorem 2.53, $\boldsymbol{H}^*_0 \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'^*_0\boldsymbol{T} \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{T}$. Consequently $\boldsymbol{H} \equiv \boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{T}$ throughout its domain $D(\boldsymbol{H}) \equiv \boldsymbol{T}^{-1}D(\boldsymbol{H}')$ as we wished to show.

I have not been able to use this theorem to characterize $\boldsymbol{H}$ intrinsically, because of my inability to determine the adjoint $\boldsymbol{H}^*_0$ of $\boldsymbol{H}_0$ either directly or indirectly. Let us, however, leave this problem for later consideration, and turn our attention to the investigation of $\boldsymbol{H}'$, $\boldsymbol{H}'_0$, and $\boldsymbol{H}_0$.

4. *The operators $\boldsymbol{H}'$ and $\boldsymbol{H}'_0$.*

THEOREM II. *Let $D(\boldsymbol{H}')$ consist of all vector functions $\boldsymbol{f} = (f_1, \cdots, f_6)$ which belong to $L_{2,6}$, and with the property that the vector*

$$
(4.1) \qquad \boldsymbol{s} = \left\{ \begin{array}{c|c} \mathbf{0} & \begin{array}{ccc} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{array} \\ \hline \begin{array}{c} \text{upper right,} \\ \text{signs changed} \end{array} & \mathbf{0} \end{array} \right\} \left\{ \begin{array}{c} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{array} \right\}
$$

also belongs to $L_{2,6}$. Let $\boldsymbol{H}'$ be the transformation which takes $\boldsymbol{f} \varepsilon D(\boldsymbol{H}')$ into $\boldsymbol{s}$. Then $\boldsymbol{H}'$ is self-adjoint.

Proof. It is evident that $D(\boldsymbol{H}')$ is a linear manifold. We prove first that $D(\boldsymbol{H}')$ is everywhere dense in $L_{2,6}$. Let D consist of all functions whose components can be expressed as linear combinations of functions which are defined as 1 inside and on an arbitrary axis-parallel parallelepiped, and zero elsewhere. Then $D \subset D(\boldsymbol{H}') \subset L_{2,6}$. But it is well known that D is dense in $L_{2,6}$. Hence also is $D(\boldsymbol{H}')$. That is, $D(\boldsymbol{H}')$ determines the closed linear manifold $L_{2,6}$. We can then prove that $\boldsymbol{H}'$ is symmetric by showing directly that the relation $(\boldsymbol{H}'\boldsymbol{f}, \boldsymbol{g}) - (\boldsymbol{f}, \boldsymbol{H}'\boldsymbol{g}) = 0$ is true for every $\boldsymbol{f}$ and $\boldsymbol{g}$ belonging to $D(\boldsymbol{H}')$. If the difference in question is written out, it will be found that the terms line up in pairs, cancelling each other, and that the desired result is obtained.

Now define $\boldsymbol{H}'_l$ by the relation $\boldsymbol{H}'_l = \boldsymbol{H}' - l\boldsymbol{I}$. Since $\boldsymbol{H}'$ is symmetric, $\boldsymbol{H}'_l$ has an inverse $\boldsymbol{H}'_l{}^{-1}$ whenever l is not real.[7] We shall prove that the domain of $\boldsymbol{H}'_l{}^{-1}$ is the entire space $L_{2,6}$. Consider the solution of the equation $\boldsymbol{H}'_l{}^{-1}\boldsymbol{f} = \boldsymbol{g}$ where $\boldsymbol{g}$ is an arbitrary element of $L_{2,6}$. In matrix form, this equation is

$$
\left\{ \begin{array}{c|c} \begin{array}{ccc} -l & 0 & 0 \\ 0 & -l & 0 \\ 0 & 0 & -l \end{array} & \begin{array}{ccc} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{array} \\ \hline \begin{array}{c} \text{upper right,} \\ \text{signs changed} \end{array} & \begin{array}{c} \text{upper} \\ \text{left} \end{array} \end{array} \right\} \left\{ \begin{array}{c} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{array} \right\} = \left\{ \begin{array}{c} g_1 \\ g_2 \\ g_3 \\ g_4 \\ g_5 \\ g_6 \end{array} \right\}.
$$

[7] Stone, Theorem 4.14.

To solve this equation for $\boldsymbol{f}$, we compute by ordinary means the inverse of $\boldsymbol{H}'_l$. The solution is then $\boldsymbol{f} = \boldsymbol{H}'_l{}^{-1}\boldsymbol{g}$, or, in matrix form,

$$
(4.2)\qquad
\begin{Bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \end{Bmatrix}
= \frac{1}{l(l^2 - x^2 - y^2 - z^2)}
\left\{\begin{array}{ccc|ccc}
x^2 - l^2 & xy & xz & 0 & -lz & ly \\
yx & y^2 - l^2 & yz & lz & 0 & -lx \\
zx & zy & z^2 - l^2 & -ly & lx & 0 \\
\hline
 & \text{upper right,} & & & \text{upper} & \\
 & \text{signs changed} & & & \text{left} &
\end{array}\right\}
\begin{Bmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \\ g_5 \\ g_6 \end{Bmatrix}
$$

It is apparent that, since l is not real, the effect of applying $\boldsymbol{H}'_l{}^{-1}$ to $\boldsymbol{g} \in L_{2,6}$ is to multiply each component of $\boldsymbol{g}$ by a bounded, measurable function. It follows that $\boldsymbol{f} \in L_{2,6}$. Now apply $\boldsymbol{H}'_l$ to both sides of (4.2). We have $\boldsymbol{H}'\boldsymbol{f} - l\boldsymbol{f} = \boldsymbol{g}$, or $\boldsymbol{H}'\boldsymbol{f} = l\boldsymbol{f} + \boldsymbol{g}$. Since the right-hand side of the last equation belongs to $L_{2,6}$, so also does the left-hand side, which is immediately identified as the vector $\boldsymbol{s}$ of (4.1). It follows that $\boldsymbol{f} \in D(\boldsymbol{H}')$. That is, $\boldsymbol{H}'_l{}^{-1}$ carries $L_{2,6}$ in a one-to-one manner into $D(\boldsymbol{H}')$, this being true for all not-real l. This is equivalent to saying that the ranges of $\boldsymbol{H}'_{+i}$ and $\boldsymbol{H}'_{-i}$ are both $L_{2,6}$. $\boldsymbol{H}'$ is consequently self-adjoint.[8]

THEOREM III. *Let $D(\boldsymbol{H}'_0)$ consist of all vector functions $\boldsymbol{f} = (f_1, \cdots, f_6)$ which belong to $L_{2,6}$, and with the property that any of their components multiplied by x, y, or z belongs to L_2. Let $\boldsymbol{H}'_0$ be the transformation which takes $\boldsymbol{f} \in D(\boldsymbol{H}'_0)$ into the vector $\boldsymbol{s}$ of* (4.1). *Then $\boldsymbol{H}'_0$ is essentially self-adjoint.*

Proof. It can be shown (i) that $D(\boldsymbol{H}'_0)$ determines the closed linear manifold $L_{2,6}$, and (ii) that $\boldsymbol{H}'_0$ is symmetric by precisely the same method as that used in Theorem II.

We now determine directly the adjoint $\boldsymbol{H}'^*_0$ of $\boldsymbol{H}'_0$. Its domain consists of those and only those elements $\boldsymbol{g} \in L_{2,6}$ such that the relation $(\boldsymbol{H}'_0\boldsymbol{f}, \boldsymbol{g}) = (\boldsymbol{f}, \boldsymbol{g}^*)$ holds for all $\boldsymbol{f} \in D(\boldsymbol{H}'_0)$ and some element $\boldsymbol{g}^* \in L_{2,6}$; and, for such an element, $\boldsymbol{H}'^*_0\boldsymbol{g} = \boldsymbol{g}^*$. Consider the equation

$$(\boldsymbol{H}'_0\boldsymbol{f}, \boldsymbol{g}) = (\boldsymbol{f}, \boldsymbol{g}^*)$$

for functions $\boldsymbol{f}$ vanishing outside an arbitrary cube. We have

$$
\begin{aligned}
&\int\!\!\int_{-A}^{A}\!\!\int [(zf_5 - yf_6)\bar{g}_1 + (-zf_4 + xf_6)\bar{g}_2 + (yf_4 - xf_5)\bar{g}_3 \\
&\quad + (-zf_2 + yf_3)\bar{g}_4 + (zf_1 - xf_3)\bar{g}_5 + (-yf_1 + xf_2)\bar{g}_6]\,dxdydz \\
&= \int\!\!\int_{-A}^{A}\!\!\int [f_1\bar{g}^*_1 + f_2\bar{g}^*_2 + f_3\bar{g}^*_3 + f_4\bar{g}^*_4 + f_5\bar{g}^*_5 + f_6\bar{g}^*_6]\,dxdydz.
\end{aligned}
$$

[8] Stone, Theorems 9.1 to 9.3.

Because of the arbitrariness of $\boldsymbol{f}$, we must have, almost everywhere,

$$
\begin{aligned}
g^*_1 &= zg_5 - yg_6 & \qquad g^*_4 &= -zg_2 + yg_3\\
g^*_2 &= -zg_4 + xg_6 & \qquad g^*_5 &= zg_1 - xg_3\\
g^*_3 &= yg_4 - xg_5 & \qquad g^*_6 &= -yg_1 + xg_2.
\end{aligned}
$$

Since these relations hold in an arbitrary cube, they must hold over all space. $\boldsymbol{H'^*}_0$ is therefore a transformation with a domain consisting of all functions $\boldsymbol{g} \varepsilon L_{2,6}$ such that $\boldsymbol{g^*}$ also belongs to $L_{2,6}$, and which takes a function in its domain into $\boldsymbol{g^*}$ as above. $\boldsymbol{H'^*}_0$ is thus identified as the transformation $\boldsymbol{H'}$, which is self-adjoint. Hence $\boldsymbol{H'}_0$ is essentially self-adjoint.

5. *The operator* $\boldsymbol{H}_0 \equiv \boldsymbol{T}^{-1}\boldsymbol{H'}_0\boldsymbol{T}$.

THEOREM IV. *Let $\boldsymbol{H}_0$ be the transformation defined by means of the relation $\boldsymbol{H}_0 \equiv \boldsymbol{T}^{-1}\boldsymbol{H'}_0\boldsymbol{T}$, where $\boldsymbol{H'}_0$ is the transformation of Theorem III. Then the domain $D(\boldsymbol{H}_0)$ of $\boldsymbol{H}_0$ consists of all vector functions $\boldsymbol{g} \varepsilon L_{2,6}$ such that the components are absolutely continuous in x, y, and z separately and have the property that their first partial derivatives with respect to x, y, or z belong to L_2. $\boldsymbol{H}_0$ is the essentially self-adjoint transformation which takes $\boldsymbol{g} \varepsilon D(\boldsymbol{H}_0)$ into the vector defined by the expression*

$$
\left\{
\begin{array}{c|c}
\begin{matrix} \\ 0 \\ \\ \end{matrix} &
\begin{matrix}
0 & -\frac{1}{i}\frac{\partial}{\partial z} & \frac{1}{i}\frac{\partial}{\partial y}\\
\frac{1}{i}\frac{\partial}{\partial z} & 0 & -\frac{1}{i}\frac{\partial}{\partial x}\\
-\frac{1}{i}\frac{\partial}{\partial y} & \frac{1}{i}\frac{\partial}{\partial x} & 0
\end{matrix}\\
\hline
\begin{matrix}\text{upper right,}\\ \text{signs changed}\end{matrix} & 0
\end{array}
\right\}
\left\{
\begin{matrix} g_1\\ g_2\\ g_3\\ g_4\\ g_5\\ g_6 \end{matrix}
\right\}
$$

Proof. Since $\boldsymbol{H'}_0$ has been shown to be essentially self-adjoint, $\boldsymbol{H}_0$ will also enjoy that property, by virtue of Stone, Theorem 2. 55. The domain of $\boldsymbol{H}_0$ is characterized as follows: if $\boldsymbol{H'}_0$ takes $\boldsymbol{f} \varepsilon D(\boldsymbol{H'}_0)$ into $\boldsymbol{f^*}$, then $\boldsymbol{H}_0$ takes $\boldsymbol{T}^{-1}\boldsymbol{f} \varepsilon D(\boldsymbol{H}_0)$ into $\boldsymbol{T}^{-1}\boldsymbol{f^*}$. If we put $\boldsymbol{g} \equiv \boldsymbol{T}^{-1}\boldsymbol{f}$, $\boldsymbol{g^*} \equiv \boldsymbol{T}^{-1}\boldsymbol{f^*} \equiv \boldsymbol{T}^{-1}[\boldsymbol{H'}_0\boldsymbol{f}]$, we must have

$$
\begin{aligned}
g^*_1 &= T^{-1}[zf_6 - yf_5] & \qquad g^*_4 &= T^{-1}[-zf_2 + yf_3]\\
g^*_2 &= T^{-1}[-zf_4 + xf_6] & \qquad g^*_5 &= T^{-1}[zf_1 - xf_3]\\
g^*_3 &= T^{-1}[yf_4 - xf_5] & \qquad g^*_6 &= T^{-1}[-yf_1 + xf_2].
\end{aligned}
$$

Because of the linearity of T^{-1}, and of the symmetry which manifests itself throughout, it will be sufficient if we put $g = T^{-1}f$, $g^* = T^{-1}[xf]$, and determine what g^* is in terms of g. We shall show first that

$$(5.1) \qquad g^*(x, y, z) = -\frac{1}{i}\ \underset{h\to 0}{\text{l.i.m.}}\ \frac{g(x+h, y, z) - g(x, y, z)}{h}.$$

We have

$$g(x, y, z) = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{(2\pi)^{3/2}} \int\int_{-A}^{A}\int e^{-i(x\xi+y\eta+z\zeta)} f(\xi, \eta, \zeta)\, d\xi d\eta d\zeta,$$

$$g^*(x, y, z) = \underset{A\to\infty}{\text{l.i.m.}}\ \frac{1}{(2\pi)^{3/2}} \int\int_{-A}^{A}\int e^{-i(x\xi+y\eta+z\zeta)} \xi f(\xi, \eta, \zeta)\, d\xi d\eta d\zeta.$$

It is well known that $|f| = |Tf| = |T^{-1}f|$. Hence

$$(5.2) \qquad \left| \frac{g(x+h, y, z) - g(x, y, z)}{ih} + g^*(x, y, z) \right| = \left| \left(\frac{e^{-ih\xi} - 1}{ih} + \xi \right) f(\xi, \eta, \zeta) \right|.$$

Now it can be shown by virtue of the mean value theorem that

$$\left| \frac{e^{-ih\xi} - 1}{ih} \right| < (2)^{1/2} |\xi|,$$

and consequently that

$$\left| \left(\frac{e^{-ih\xi} - 1}{ih} + \xi \right) f(\xi, \eta, \zeta) \right| < k\, |\xi f(\xi, \eta, \zeta)|,$$

where k is a suitable constant. In addition,

$$\lim_{h\to 0} \left(\frac{e^{-ih\xi} - 1}{ih} + \xi \right) = 0.$$

These last two relations are sufficient, by a familiar theorem regarding passage to the limit under the sign of integration, to insure the convergence of the right-hand side of (5.2) to zero as $h \to 0$. This in turn implies (5.1), as we wished to show.

We propose now to show that

$$g^*(x, y, z) = -\frac{1}{i}\ \frac{\partial}{\partial x} g(x, y, z)$$

almost everywhere. Since g and g^* belong to L_2, we have by (5.1)

$$\lim_{h\to 0}\int_a^x\int_b^y\int_c^z \frac{g(\xi+h,\eta,\zeta)-g(\xi,\eta,\zeta)}{ih}\,d\xi d\eta d\zeta$$
$$=-\int_a^x\int_b^y\int_c^z g^*(\xi,\eta,\zeta)\,d\xi d\eta d\zeta.$$

That is,

$$\lim_{h\to 0}\frac{1}{h}\left[\int_x^{x+h}-\int_a^{a+h}\right]\int_b^y\int_c^z g(\xi,\eta,\zeta)\,d\xi d\eta d\zeta$$
$$=-i\int_a^x\int_b^y\int_c^z g^*(\xi,\eta,\zeta)\,d\xi d\eta d\zeta,$$

and

$$(5.3)\quad \int_b^y\int_c^z [g(x,\eta,\zeta)-g(a,\eta,\zeta)]\,d\eta d\zeta$$
$$=-i\int_a^x\int_b^y\int_c^z g^*(\xi,\eta,\zeta)\,d\xi d\eta d\zeta$$

for all x outside a one-dimensional null set dependent on y and z. To show that this relation holds for all x, y, z outside a three-dimensional null set, we argue as follows. Since the relation is one between measurable functions, the set on which it fails to hold is certainly a measurable set. Denote this three-dimensional set by S. The two-dimensional intersection S_z of S with $z=$ constant is measurable for all z outside a one-dimensional null set. For all such z, the one-dimensional intersection S_{zy} of S_z with $y=$ constant is measurable for all y outside a one-dimensional null set. But for all such y, the projection of S_{zy} on the axis of x is of measure zero, since (5.3) holds for all x outside a one-dimensional null set dependent on y and z. By a theorem of Fubini,[9] it follows that the two-dimensional set S_z is of measure zero. We need only repeat the argument to show that S is likewise of measure zero.

Now change the order of integration on the right-hand side of (5.3), and differentiate with respect to y. Then

$$\int_c^z [g(x,y,\zeta)-g(a,y,\zeta)]\,d\zeta=-i\int_c^z\int_a^x g^*(\xi,y,\zeta)\,d\xi d\zeta$$

for all y outside a one-dimensional null set dependent on x and z. We have only to repeat the argument used above to show that this relation holds for all x, y, z outside a three-dimensional null set. Now differentiate with respect to z.

$$g(x,y,z)-g(a,y,z)=-i\int_a^x g^*(\xi,y,z)\,d\xi$$

[9] See, e. g., C. Caratheodory, *Vorlesungen über Reele Funktionen* (1918), Satz 3, p. 628.

for all z outside a one-dimensional null set dependent on x and y. We can likewise show that this relation holds except in a three-dimensional null set. It is now evident that for a fixed y and z outside a certain two-dimensional null set, $g(x, y, z)$ is equal for almost all x to an absolutely continuous function of x, whose derivative is $-ig^*(x, y, z)$ for almost all x. A third and final repetition of the argument used above results in the conclusion that

$$g^*(x, y, z) = -\frac{1}{i} \frac{\partial}{\partial x} g(x, y, z)$$

almost everywhere.

It remains to show conversely that, if we put as before $f = Tg$, $f^* = Tg^*$, then $f^* = xf$. Let

$$f(x, y, z; A) = \frac{1}{(2\pi)^{3/2}} \int\int_{-A}^{A}\int e^{i(x\xi+y\eta+z\zeta)} g(\xi, \eta, \xi)\, d\xi d\eta d\zeta,$$

$$f^*(x, y, z; A) = \frac{-1}{i(2\pi)^{3/2}} \int\int_{-A}^{A}\int e^{i(x\xi+y\eta+z\zeta)} \frac{\partial}{\partial \xi} g(\xi, \eta, \zeta)\, d\xi d\eta d\zeta.$$

Then $f(x, y, z; A)$ and $f^*(x, y, z; A)$ converge in the mean as $A \to \infty$ to $f(x, y, z)$ and $f^*(x, y, z)$ respectively. We can, as well as not, assume that A runs over the positive integers. It follows that there exists a subsequence of integers, say $\{m\}$, for which, almost everywhere, $f(x, y, z; m)$ and $f^*(x, y, z; m)$ converge in the ordinary sense to $f(x, y, z)$ and $f^*(x, y, z)$ respectively. If now we define the function $h(x, y, z)$ by the relation

$$h(x, y, z) \equiv f^*(x, y, z) - xf(x, y, z),$$

then the function

$$h(x, y, z; m) \equiv f^*(x, y, z; m) - xf(x, y, z; m)$$

tends almost everywhere to $h(x, y, z)$ as $m \to \infty$. In addition, $h(x, y, z)$ is integrable over every finite interval. We propose to show that $h(x, y, z) \equiv 0$ almost everywhere. If we integrate $f^*(x, y, z; m)$ by parts with respect to ξ, we have

$$\begin{aligned} h(x, y, z; m) &= \frac{-1}{i(2\pi)^{3/2}} \int_{-m}^{m}\int \left[e^{i(x\xi+y\eta+z\zeta)} g(\xi, \eta, \zeta)\right]_{\xi=-m}^{\xi=m} d\eta d\zeta \\ &= -\frac{e^{imx}}{i(2\pi)^{3/2}} \int_{-m}^{m}\int e^{i(y\eta+z\zeta)} g(m, \eta, \zeta)\, d\eta d\zeta \\ &\quad + \frac{e^{-imx}}{i(2\pi)^{3/2}} \int_{-m}^{m}\int e^{i(y\eta+z\zeta)} g(-m, n, \zeta)\, d\eta d\zeta. \end{aligned}$$

That is, we can write $h(x, y, z; m)$ in the form

$$h(x, y, z; m) = \phi(y, z; m) e^{imx} + \psi(y, z; m) e^{-imx}.$$

Now separate all quantities into their real and imaginary parts; i. e., put $h = h_1 + ih_2$, $\phi = \phi_1 + i\phi_2$, etc. The above relation then becomes

$$(5.4) \quad \begin{aligned} h_1(x, y, z; m) &= (\phi_1 + \psi_1) \cos mx + (-\phi_2 + \psi_2) \sin mx, \\ h_2(x, y, z; m) &= (\phi_2 + \psi_2) \cos mx + (\phi_1 - \psi_1) \sin mx, \end{aligned}$$

and in addition,

$$(5.5) \qquad \lim_{m\to\infty} h_j(x, y, z; m) = h_j(x, y, z), \qquad (j = 1, 2),$$

almost everywhere.

If now we fix y and z, ϕ and ψ become functions of m alone, and the right-hand side of either of the relations of (5.4) can be written in the form $B_m \cos mx + C_m \sin mx$, where the coefficients B_m and C_m are real numbers dependent on m. By a theorem of Steinhaus,[10]

$$\overline{\lim_{m\to\infty}} \, | B_m \cos mx + C_m \sin mx | = \overline{\lim_{m\to\infty}} \, (B_m{}^2 + C_m{}^2)^{1/2}$$

for all x outside a null set. The left-hand side is finite for all x outside a null set (since h_1 and h_2 are integrable over every finite interval), and the right-hand side is independent of x. Hence B_m and C_m remain bounded. Now write

$$B_m \cos mx + C_m \sin mx = (B_m{}^2 + C_m{}^2)^{1/2} \cos m(x - \alpha_m),$$

where $\alpha_m = \tan^{-1} C_m/B_m$; and note that for almost all x,

$$\overline{\lim_{m\to\infty}} \cos m(x - \alpha_m) = +1, \qquad \underline{\lim_{m\to\infty}} \cos m(x - \alpha_m) = -1.$$

Furthermore, we can pick out a subsequence, say $\{\mu\}$ of m's such that $(B_\mu{}^2 + C_\mu{}^2)^{1/2}$ tends to its upper limit. In that case,

$$\overline{\lim_{\mu\to\infty}} \, (B_\mu{}^2 + C_\mu{}^2)^{1/2} \cos \mu(x - \alpha_\mu) = +\overline{\lim_{\mu\to\infty}} \, (B_\mu{}^2 + C_\mu{}^2)^{1/2} \geq 0,$$

$$\underline{\lim_{\mu\to\infty}} \, (B_\mu{}^2 + C_\mu{}^2)^{1/2} \cos \mu(x - \alpha_\mu) = -\overline{\lim_{\mu\to\infty}} \, (B_\mu{}^2 + C_\mu{}^2)^{1/2} \leq 0.$$

But the limits on the left are identical by (5.5). Consequently $h(x, y, z) \equiv 0$ almost everywhere, as we wished to show. This completes the proof of Theorem IV.

[10] *Wiadomosci Matematyczne,* vol. 24 (1920), pp. 197-201.

Let us consider what has been accomplished up to this point. We have defined the operators $\boldsymbol{H}'$, $\boldsymbol{H}'_0$, and $\boldsymbol{H}_0$; and have shown $\boldsymbol{H}'$ self-adjoint, and $\boldsymbol{H}'_0$ and $\boldsymbol{H}_0$ essentially self-adjoint. Were we able to determine the adjoint $\boldsymbol{H}^*_0$ of $\boldsymbol{H}_0$ we would have, as proved in Theorem I (Section 3), a self-adjoint transformation identical with $\boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{T}$ throughout its domain $\boldsymbol{T}^{-1}D(\boldsymbol{H}')$. As I remarked previously, I have been unable to determine $\boldsymbol{H}^*_0$ either directly or indirectly. For all practical purposes, however, the exact determination of $\boldsymbol{H}^*_0$ is unnecessary, since the functions to which one might have occasion to apply the theory would probably satisfy much more restrictive conditions than those required of functions in the domain of $\boldsymbol{H}^*_0$. Indeed, most of the functions considered in classical electrodynamics would be included in the domain of $\boldsymbol{H}_0$ (i. e., possess the necessary derivatives, belong to $L_{2,6}$, etc.). Consequently, since we are sure $\boldsymbol{H}^*_0$ exists, we shall hereafter refer to it as the self-adjoint transformation $\boldsymbol{H}$ with domain $D(\boldsymbol{H})$.

6. *The operator $\boldsymbol{E}'(\lambda)$ corresponding to $\boldsymbol{H}'$.* In the calculation of the operator $\boldsymbol{E}'(\lambda)$ we make use of Stone, Theorem 5. 10, to wit: If $\boldsymbol{H}'$ is a given self-adjoint transformation, the corresponding "resolution of the identity" $\boldsymbol{E}'(\lambda)$ can be determined from the relation

$$
\begin{aligned}
\tfrac{1}{2}\{[(\boldsymbol{E}'(\mu)\boldsymbol{f},\boldsymbol{g}) + (\boldsymbol{E}'(\mu-0)\boldsymbol{f},\boldsymbol{g})] \\
-[(\boldsymbol{E}'(\nu)\boldsymbol{f},\boldsymbol{g}) + (\boldsymbol{E}'(\nu-0)\boldsymbol{f},\boldsymbol{g})]\} \\
= -\frac{1}{2\pi i}\lim_{\epsilon\to 0}\int_C (\boldsymbol{H}'_l{}^{-1}\boldsymbol{f},\boldsymbol{g})\,dl,
\end{aligned}
\tag{6.1}
$$

where $\boldsymbol{f}$ and $\boldsymbol{g}$ are arbitrary elements of $L_{2,6}$, and where the contour over which the integral is taken consists of two oriented polygonal lines whose vertices, in order, are $\mu+i\epsilon$, $\mu+i\alpha$, $\nu+i\alpha$, $\nu+i\epsilon$, and $\nu-i\epsilon$, $\nu-i\alpha$, $\mu-i\alpha$, $\mu-i\epsilon$, respectively; the real numbers μ, ν, α, ϵ, being subject to the inequalities $\nu<\mu$, $0<\epsilon<\alpha$.

To obtain $\boldsymbol{E}'(\lambda)$ from (6.1) we put $\mu=\lambda+\delta$, $\delta>0$, and allow δ to tend to zero, and ν to tend to $-\infty$. For by the properties[11] of $\boldsymbol{E}'(\lambda)$,

$$
\begin{gathered}
\lim_{\delta\to 0}(\boldsymbol{E}'(\lambda+\delta)\boldsymbol{f},\boldsymbol{g}) = (\boldsymbol{E}'(\lambda+0)\boldsymbol{f},\boldsymbol{g}) = (\boldsymbol{E}'(\lambda)\boldsymbol{f},\boldsymbol{g}),\\
\lim_{\delta\to 0}(\boldsymbol{E}'(\lambda+\delta-0)\boldsymbol{f},\boldsymbol{g}) = (\boldsymbol{E}'(\lambda+0)\boldsymbol{f},\boldsymbol{g}) = (\boldsymbol{E}'(\lambda)\boldsymbol{f},\boldsymbol{g}),\\
\lim_{\nu\to-\infty}(\boldsymbol{E}'(\nu)\boldsymbol{f},\boldsymbol{g}) = 0, \qquad \lim_{\nu\to-\infty}(\boldsymbol{E}'(\nu-0)\boldsymbol{f},\boldsymbol{g}) = 0,
\end{gathered}
$$

and the left-hand side of (6.1) becomes simply $(\boldsymbol{E}'(\lambda)\boldsymbol{f},\boldsymbol{g})$. It is to be noted that the domain of $\boldsymbol{E}'(\lambda)$ is the entire space $L_{2,6}$.

[11] Stone, Definition 5. 1.

For convenience of notation, let us write $(\boldsymbol{E}'(\mu,\nu)\boldsymbol{f},\boldsymbol{g})$ for the left-hand side of (6.1), and put $\boldsymbol{R}\equiv\boldsymbol{H}'_l{}^{-1}$ (see relation (4.2) for the latter). Denote their elements by $E'_{jk}(\mu,\nu)$ and R_{jk} respectively. Then (6.1) becomes

$$(6.2)\qquad \iiint_{-\infty}^{+\infty} \sum_{j,k=1}^{6} E'_{jk}(\mu,\nu) f_k \bar{g}_j \, dxdydz = -\frac{1}{2\pi i}\lim_{\epsilon\to 0}\int_C \left[\iiint_{-\infty}^{+\infty}\sum_{j,k=1}^{6} R_{jk} f_k \bar{g}_j \, dxdydz\right] dl.$$

The change of order of integration on the right-hand side of this equation can easily be justified because of the simple way in which l enters into the expression for the R_{jk}. The right-hand member can thus be written

$$(6.3)\qquad \lim_{\epsilon\to 0}\iiint_{-\infty}^{+\infty}\sum_{j,k=1}^{6}\left[-\frac{1}{2\pi i}\int_C R_{jk}dl\right] f_k\bar{g}_j \, dxdydz.$$

We proceed now to calculate the various contour integrals involved. Inspection of the matrix $\boldsymbol{R}$ shows that the calculation requires integrating the following three fractions, or combinations thereof, where for simplicity we put $r\equiv(x^2+y^2+z^2)^{1/2}$: $1/(l^2-r^2)$, $l/(l^2-r^2)$, $1/l(l^2-r^2)$. These in turn break down into partial fractions whose contour integrals are easily found. The following results are obtained:

$$\frac{1}{2\pi i}\int_C \frac{dl}{l^2-r^2} = \frac{1}{2\pi r}\left[\tan^{-1}\frac{\nu-r}{\epsilon} - \tan^{-1}\frac{\mu-r}{\epsilon} - \tan^{-1}\frac{\nu+r}{\epsilon} + \tan^{-1}\frac{\mu+r}{\epsilon}\right]$$

$$\frac{1}{2\pi i}\int_C \frac{ldl}{l^2-r^2} = \frac{1}{2\pi}\left[\tan^{-1}\frac{\nu-r}{\epsilon} - \tan^{-1}\frac{\mu-r}{\epsilon} + \tan^{-1}\frac{\nu+r}{\epsilon} - \tan^{-1}\frac{\mu+r}{\epsilon}\right]$$

$$\frac{1}{2\pi i}\int_C \frac{dl}{l(l^2-r^2)} = \frac{1}{2\pi r^2}\left[\tan^{-1}\frac{\nu-r}{\epsilon} - \tan^{-1}\frac{\mu-r}{\epsilon} + \tan^{-1}\frac{\nu+r}{\epsilon} - \tan^{-1}\frac{\mu+r}{\epsilon} - 2\tan^{-1}\frac{\nu}{\epsilon} + 2\tan^{-1}\frac{\mu}{\epsilon}\right].$$

Note now that as $\epsilon\to 0$ through positive values, these expressions converge boundedly to their respective limits. We are consequently justified in passing to the limit under the sign of integration in (6.3). That is, if we put

$$P_{jk}\equiv -\frac{1}{2\pi i}\lim_{\epsilon\to 0}\int_C R_{jk}dl,$$

we can conclude that

$$\iiint_{-\infty}^{+\infty}\sum_{j,k=1}^{6} E'_{jk}(\mu,\nu) f_k\bar{g}_j \, dxdydz = \iiint_{-\infty}^{+\infty}\sum_{j,k=1}^{6} P_{jk} f_k\bar{g}_j \, dxdydz.$$

Finally, in view of the fact that f and g are arbitrary elements of $L_{2,6}$, we must have $E'_{jk}(\mu,\nu)f_k = P_{jk}f_k$ almost everywhere.

In calculating the limits of the contour integrals, there are three cases to be considered, namely, (i) $0 < \nu < \mu$, (ii) $\nu < 0 < \mu$, (iii) $\nu < \mu < 0$. (Recall that μ was taken greater than ν). But since, in passing from $E'_{jk}(\mu,\nu)$ to $E'_{jk}(\lambda)$, we are going to let ν tend to $-\infty$, we may as well take ν negative, which throws out case (i). Also, in case (ii), we shall assume $\mu < -\nu$. If, then, we put $\mu = \lambda + \delta$, $\delta > 0$, and allow δ to tend to zero, ν to tend to $-\infty$, $E'_{jk}(\mu,\nu)$ tends in the mean to $E'_{jk}(\lambda)$, and we obtain the following matrices:

$$\underset{(\lambda<0)}{\boldsymbol{E}'(\lambda)} = \begin{cases} \boldsymbol{O} \qquad \text{(null matrix)}, & r \leq -\lambda \\ \dfrac{-1}{2r^2}\left\{\begin{array}{ccc|ccc} x^2-r^2 & xy & xz & 0 & zr & -yr \\ yx & y^2-r^2 & yz & -zr & 0 & xr \\ zx & zy & z^2-r^2 & yr & -xr & 0 \\ \hline \multicolumn{3}{c|}{\text{upper right, signs changed}} & \multicolumn{3}{c}{\text{upper left}} \end{array}\right\}, & r > -\lambda \end{cases}$$

$$\underset{(\lambda>0)}{\boldsymbol{E}'(\lambda)} = \begin{cases} \boldsymbol{I} \qquad \text{(identity matrix)}, & r \leq \lambda \\ \dfrac{1}{2r^2}\left\{\begin{array}{ccc|ccc} x^2+r^2 & xy & xz & 0 & -zr & yr \\ yx & y^2+r^2 & yz & zr & 0 & -xr \\ zx & zy & z^2+r^2 & -yr & xr & 0 \\ \hline \multicolumn{3}{c|}{\text{upper right, signs changed}} & \multicolumn{3}{c}{\text{upper left}} \end{array}\right\}, & r > \lambda. \end{cases}$$

The operators which for the most part we shall have occasion to use, however, are not those just calculated, but rather the operators $[\boldsymbol{E}'(\lambda) - \boldsymbol{E}'(\mu)]$, as they appear in the cases $\mu < \lambda < 0$ and $0 < \mu < \lambda$. If we choose the following nine elements as basic:

$$E'_{11} = \frac{r^2-x^2}{2r^2} \qquad E'_{12} = \frac{-xy}{2r^2} \qquad E'_{35} = \frac{-x}{2r}$$

$$E'_{22} = \frac{r^2-y^2}{2r^2} \qquad E'_{13} = \frac{-xz}{2r^2} \qquad E'_{16} = \frac{-y}{2r}$$

$$E'_{33} = \frac{r^2-z^2}{2r^2} \qquad E'_{23} = \frac{-yz}{2r^2} \qquad E'_{24} = \frac{-z}{2r};$$

we can express the required $[\boldsymbol{E}'(\lambda) - \boldsymbol{E}'(\mu)]$ in terms of them as follows:

$$[\boldsymbol{E}'(\lambda)-\boldsymbol{E}'(\mu)] \doteq \begin{cases} \left\{\begin{array}{ccc|ccc} E'_{11} & E'_{12} & E'_{13} & 0 & E'_{24} & -E'_{16} \\ E'_{12} & E'_{22} & E'_{23} & -E'_{24} & 0 & E'_{35} \\ E'_{13} & E'_{23} & E'_{33} & E'_{16} & -E'_{35} & 0 \\ \hline \multicolumn{3}{c|}{\text{upper right, signs changed}} & \multicolumn{3}{c}{\text{upper left}} \end{array}\right\}, & -\lambda < r \leq -\mu \\ \boldsymbol{O}, \text{ elsewhere.} & \end{cases}$$
$$(\mu < \lambda < 0)$$

$$[\boldsymbol{E}'(\lambda)-\boldsymbol{E}'(\mu)] = \begin{cases} \left\{\begin{array}{ccc|ccc} E'_{11} & E'_{12} & E'_{13} & 0 & -E'_{24} & E'_{16} \\ E'_{12} & E'_{22} & E'_{23} & E'_{24} & 0 & -E'_{25} \\ E'_{13} & E'_{23} & E'_{33} & -E'_{16} & E'_{35} & 0 \\ \hline \multicolumn{3}{c|}{\text{upper right, signs changed}} & \multicolumn{3}{c}{\text{upper left}} \end{array}\right\}, & \mu < r \leq \lambda. \\ \boldsymbol{O}, \text{ elsewhere.} & \end{cases}$$
$$(0 < \mu < \lambda)$$

In addition to these, we shall have occasion to use the matrix $[\boldsymbol{E}'(0+0) - \boldsymbol{E}'(0-0)]$. Note that $\boldsymbol{E}'(0+0)$ can be had as a limiting case of $\boldsymbol{E}'(\lambda)$, $\lambda > 0$, $\lambda \to 0$. Thus, $\boldsymbol{E}'(0+0) = \boldsymbol{E}'(\lambda)$ for $\lambda > 0$, $r > \lambda$; and similarly, $\boldsymbol{E}'(0-0) = \boldsymbol{E}'(\lambda)$ for $\lambda < 0$, $r > -\lambda$. The matrix in question can then easily be shown to be

$$[\boldsymbol{E}'(0+0)-\boldsymbol{E}'(0-0)] = \frac{1}{r^2}\left\{\begin{array}{ccc|c} x^2 & xy & xz & \\ yx & y^2 & yz & 0 \\ zx & zy & z^2 & \\ \hline \multicolumn{3}{c|}{0} & \text{upper left} \end{array}\right\}, \quad 0 < r \leq +\infty.$$

7. *Characteristic values and elements of* $\boldsymbol{H}'$ *and* $\boldsymbol{H}$. It is convenient at this point to say a few words regarding characteristic values and characteristic elements. For through a discussion of them, we are not only provided with a partial check on the calculation of $\boldsymbol{E}'(\lambda)$, but we shall in addition obtain a result which will be of use later on. A characteristic value of $\boldsymbol{H}'$ is defined [12] as a value of l for which $\boldsymbol{H}'_l$ has no inverse. Now the determinant of the matrix $\boldsymbol{H}'_l$ is $l^2(l^2-r^2)^2$; hence $l=0$ is the only characteristic value of $\boldsymbol{H}'$. $l = \pm r$ are not characteristic values, since for a fixed l, this equality holds only over the surface of a sphere—i. e., over a three-dimensional set of measure zero. A characteristic element of $\boldsymbol{H}'$ is defined as an element $\boldsymbol{g} \,\varepsilon\, L_{2,6}$,

[12] Stone, Definition 4. 2.

$g \neq 0$, such that $H'g = lg$, where l is a characteristic value. On the other hand, a necessary and sufficient condition [13] that g be a characteristic element of H' corresponding to the characteristic value l is that

$$[E'(l+0) - E'(l-0)]g = g.$$

If, then, the expressions obtained for $E'(\lambda)$ are correct, the solutions of these last two equations for the characteristic value $l = 0$ must be the same.

The matrix equation $[E'(0+0) - E'(0-0)]g = g$ leads to six scalar equations, of which the first three are

$$\begin{aligned} (x^2 - r^2)g_1 + \quad xyg_2 + \quad xzg_3 &= 0 \\ yxg_1 + (y^2 - r^2)g_2 + \quad yzg_3 &= 0 \\ zxg_1 + \quad zyg_2 + (z^2 - r^2)g_3 &= 0, \end{aligned}$$

and the second three are different only in that g_1, g_2, g_3 are replaced by g_4, g_5, g_6 respectively. The two sets of three equations are independent; and in either case, for a fixed x, y, z, the determinant of the matrix of coefficients is zero, and the rank of the matrix is 2. Hence, in either case, any one of the three components involved may be chosen arbitrarily, and the rest will be uniquely determined. Put $g_1 = xp_1$, $g_4 = xp_2$, where p_1 and p_2 are to a certain extent arbitrary measurable functions of x, y, and z. Then

$$g = (xp_1, yp_1, zp_1, xp_2, yp_2, zp_2).$$

Our choice of p_1 and p_2 is restricted in that the resulting vector g must belong to $L_{2,6}$. It is easy to verify that the solution of the equation $H'g = 0$ has the same form, as was to be expected.

We have just shown that a characteristic element of H' must be of the form $g = (xp_1, yp_1, zp_1, xp_2, yp_2, zp_2)$, where p_1 and p_2 are any measurable functions such that $g \varepsilon L_{2,6}$. Conversely, any such functions p_1 and p_2 will yield a g in the domain of H' such that $H'g = 0$. The set P' of all such functions g is a closed linear manifold, called the characteristic manifold of H'. If the characteristic manifold of H be denoted by P, then, since the characteristic values of H and H' are identical [14] we must have $P \equiv T^{-1}P'$. Suppose now that we restrict ourselves to a dense linear manifold P'_0 of P', and calculate $T^{-1}P'_0$. We shall obtain a linear manifold P_0 which will be dense in P. This follows immediately from the fact that distances are preserved under the Fourier transformation.

[13] Stone, Theorem 5.13.

[14] Stone, Theorem 4.3.

Let us take as P'_0 the set of all functions $\boldsymbol{g} \varepsilon L_{2,6}$ of the form

$$\boldsymbol{g} = (xp_1, yp_1, zp_1, xp_2, yp_2, zp_2),$$

where p_1 and p_2 belong to L_2. It is not difficult to show that this set is dense in P'. If now we apply $\boldsymbol{T}^{-1}$ to $\boldsymbol{g}$, we obtain as P_0 the set of all functions of the form

$$\boldsymbol{f} = \left(\frac{\partial p_1}{\partial x}, \frac{\partial p_1}{\partial y}, \frac{\partial p_1}{\partial z}, \frac{\partial p_2}{\partial x}, \frac{\partial p_2}{\partial y}, \frac{\partial p_2}{\partial z}\right),$$

where p_1 and p_2 belong to L_2, are absolutely continuous in x, y, and z separately, and have the property that their first partial derivatives with respect to x, y, or z belong to L_2. This result follows directly from our work in Theorem IV.

8. *The operator* $\boldsymbol{E}(\lambda)$ *corresponding to* $\boldsymbol{H}$. It has already been pointed out in Section 2 that $\boldsymbol{Hf} = \boldsymbol{T}^{-1}\boldsymbol{H}'\boldsymbol{Tf}$ implies $\boldsymbol{E}(\lambda)\boldsymbol{f} = \boldsymbol{T}^{-1}\boldsymbol{E}'(\lambda)\boldsymbol{Tf}$. Since this relation holds for an arbitrary element $\boldsymbol{f} \varepsilon L_{2,6}$ we are entitled to write $E_{jk}(\lambda)f_k = T^{-1}E'_{jk}(\lambda)Tf_k$. A difficulty arises in the calculation of the $E_{jk}(\lambda)$ due to the fact that, in most instances, the $E'_{jk}(\lambda)$, as functions of x, y, z, do not belong to L_2, as we shall presently require. We shall find ultimately that it suits our purposes just as well to calculate

$$(8.1) \qquad [E_{jk}(\lambda) - E_{jk}(\mu)]f_k = T^{-1}[E'_{jk}(\lambda) - E'_{jk}(\mu)]Tf_k$$

for the cases $\mu < \lambda < 0$, and $0 < \mu < \lambda$. Put

$$[E_{jk}(\lambda) - E_{jk}(\mu)] \equiv E_{jk}(\Delta) \equiv E_{jk}(x, y, z; \Delta).$$

Then (8.1) becomes

$$(8.2) \quad E_{jk}(x, y, z; \Delta)f_k(x, y, z)$$
$$= \frac{1}{(2\pi)^{3/2}} \int\int_{-\infty}^{+\infty}\int e^{-i(xx_2+yy_2+zz_2)} E'_{jk}(x_2, y_2, z_2; \Delta)\; Tf_k(x_2, y_2, z_2)\, dx_2 dy_2 dz_2.$$

Let us examine the right-hand member of this equation. It exists as a function of x, y, z, and belongs to L_2. To establish this fact, we argue as follows: since $f_k \varepsilon L_2$, so also does Tf_k. The effect of applying $E'_{jk}(\Delta)$ to any function is to multiply it by a bounded measurable function or by zero according as the point (x, y, z) lies inside or outside a region bounded by two concentric spherical surfaces. Hence $E'_{jk}(\Delta)Tf_k \varepsilon L_2$, and vanishes outside the above mentioned region. The right-hand member of (8.2) is nothing

but the inverse Fourier transformation applied to $E'_{jk}(\Delta)Tf_k$ and as such, it exists and belongs to L_2. Now let us look at the integrand from another point of view. Put

$$g_{jk}(x_2, y_2, z_2; x, y, z; \Delta) = e^{-i(xx_2+yy_2+zz_2)}E'_{jk}(x_2, y_2, z_2; \Delta).$$

Then g_{jk}, as a function of x_2, y_2, z_2, belongs to L_2. But it is well known that if f and $g \varepsilon L_2$, then

$$\int\int_{-\infty}^{+\infty}\int g \cdot Tf = \int\int_{-\infty}^{+\infty}\int Tg \cdot f,$$

both integrals being absolutely convergent. Thus, if we put for convenience $\xi = x_1 - x$, $\eta = y_1 - y$, $\zeta = z_1 - z$, we are entitled to write, instead of (8. 2), the following:

$$\text{(8 3)} \quad \begin{aligned} &E_{jk}(x, y, z; \Delta)f_k(x, y, z) \\ &\quad = \frac{1}{(2\pi)^{3/2}}\int\int_{-\infty}^{+\infty}\int G_{jk}(\xi, \eta, \zeta; \Delta)f_k(x_1, y_1, z_1)\,dx_1dy_1dz_1, \\ &G_{jk}(\xi, \eta, \zeta; \Delta) \\ &\quad = \frac{1}{(2\pi)^{3/2}}\int\int_{-\infty}^{+\infty}\int e^{i(\xi x_2+\eta y_2+\zeta z_2)}E'_{jk}(x_2, y_2, z_2; \Delta)\,dx_2dy_2dz_2. \end{aligned}$$

We shall leave the $E_{jk}(\Delta)f_k$ in the integral form as above, and proceed to calculate the G_{jk}. Because of the fact that the $E'_{jk}(\Delta)$ vanish outside a region between two spherical surfaces, it is advisable to work with spherical rather than with rectangular coördinates. To this end, we put $x_2 = r\cos\theta\sin\phi$, $y_2 = r\sin\theta\sin\phi$, $z_2 = r\cos\phi$. We note next that the $E'_{jk}(\Delta)$ are made up of ten fractions, or combinations thereof. They are $1/2$; x^2, y^2, z^2, xy, xz, and yz, each divided by $2r^2$; and x, y, and z, each divided by $2r$. That is, if we put for convenience $\omega \equiv \xi\cos\theta\sin\phi + \eta\sin\theta\sin\phi + \zeta\cos\phi$, we must evaluate the following ten integrals:

$$I_1 = \int_\mu^\lambda\int_0^{2\pi}\int_0^\pi e^{ir\omega}\, r^2\sin\phi\; drd\theta d\phi$$

$$I_2 = \int_\mu^\lambda\int_0^{2\pi}\int_0^\pi e^{ir\omega}\, r^2\sin^3\phi\cos^2\theta\; drd\theta d\phi$$

$$I_3 = \int_\mu^\lambda\int_0^{2\pi}\int_0^\pi e^{ir\omega}\, r^2\sin^3\phi\sin^2\theta\; drd\theta d\phi$$

$$I_4 = \int_\mu^\lambda\int_0^{2\pi}\int_0^\pi e^{ir\omega}\, r^2\sin\phi\cos^2\phi\; drd\theta d\phi$$

$$I_5 = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin^3\phi \sin\theta \cos\theta \, dr d\theta d\phi$$

$$I_6 = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin^2\phi \cos\phi \cos\theta \, dr d\theta d\phi$$

$$I_7 = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin^2\phi \cos\phi \sin\theta \, dr d\theta d\phi$$

$$I_8 = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin^2\phi \cos\theta \, dr d\theta d\phi$$

$$I_9 = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin^2\phi \sin\theta \, dr d\theta d\phi$$

$$I_{10} = \int_\mu^\lambda \int_0^{2\pi} \int_0^\pi e^{ir\omega}\, r^2 \sin\phi \cos\phi \, dr d\theta d\phi.$$

These integrals, as written, are for the case $0 < \mu < \lambda$. Those for the case $\mu < \lambda < 0$ differ only in that the lower and upper limits for r are, respectively, $-\lambda$ and $-\mu$. We shall not go into any great detail regarding the calculation of these integrals, but it is perhaps advisable to note the various devices used. Let us first consider the integration with respect to θ of I_1. This involves

$$\int_0^{2\pi} e^{ir(b \sin\theta + c\cos\theta)}\, d\theta,$$

where we have put $b \equiv \eta \sin\phi$, $c \equiv \xi \sin\phi$. If now we let $\theta = \psi + \delta$ where δ is defined by the relation $\delta = \tan^{-1} b/c$, and if we put $d \equiv r(b^2 + c^2)^{1/2}$, the above integral becomes

$$\int_{-\delta}^{2\pi-\delta} e^{id\cos\psi}\, d\psi = \int_0^{2\pi} e^{id\cos\psi}\, d\psi = 2\int_0^\pi e^{id\cos\psi}\, d\psi = 2\pi J_0(d),$$

where J_0 is the Bessel function of the first kind and 0-th order. Note that, because of our substitutions, $J_0(d) \equiv J_0(r\sqrt{\xi^2+\eta^2}\sin\phi)$.

Similarly, I_2 involves

$$\int_0^{2\pi} e^{ir(b\sin\theta + c\cos\theta)} \cos^2\theta d\theta = -\int_0^{2\pi} \frac{1}{r^2}\frac{\partial^2}{\partial c^2}\left[e^{ir(b\sin\theta + c\cos\theta)}\right] d\theta$$
$$= -\frac{1}{r^2}\frac{\partial^2}{\partial c^2}\left[\int_0^{2\pi} e^{ir(b\sin\theta + c\cos\theta)}\, d\theta\right] = -\frac{2\pi}{r^2}\frac{\partial^2}{\partial c^2}\left[J_0(d)\right].$$

This last term can be broken down as follows:

$$\frac{\partial^2}{\partial c^2}\left[J_0(d)\right] = r^2 \frac{\xi^2}{\xi^2+\eta^2} J_0''(d) + \frac{r}{\sin\phi}\,\frac{\eta^2}{(\xi^2+\eta^2)^{3/2}} J'_0(d),$$

where the prime denotes differentiation with respect to the argument d. We have also $J_0'(d) = -J_1(d)$, $J_0''(d) = \frac{1}{2}[J_2(d) - J_0(d)]$. Hence I_2 can be

written as the sum of three integrals, involving J_0, J_1, and J_2 respectively. The integrals I_3 to I_{10} are treated in a similar manner.

The integration with respect to ϕ amounts to finding expressions for the following:

$$\text{(A)} \qquad \int_0^\pi e^{ir\zeta\cos\phi} J_n(r\sqrt{\xi^2+\eta^2}\sin\phi)\sin^{n+1}\phi\, d\phi$$

$$\text{(B)} \qquad \int_0^\pi e^{ir\zeta\cos\phi} J_n(r\sqrt{\xi^2+\eta^2}\sin\phi)\cos\phi\sin^{n+1}\phi\, d\phi$$

$$\text{(C)} \qquad \int_0^\pi e^{ir\zeta\cos\phi} J_n(r\sqrt{\xi^2+\eta^2}\sin\phi)\cos^2\phi\sin^{n+1}\phi\, d\phi.$$

The following formula, which has been derived by both N. Sonine and N. Nielsen,[15] is of basic importance in this respect:

$$\text{(S)} \qquad \int_0^{\pi/2} J_m(aq\cos\phi)J_n(az\sin\phi)\cos^{m+1}\phi\sin^{n+1}\phi\, d\phi = \frac{q^m z^n}{a}\,\frac{J_{m+n+1}(a\sqrt{q^2+z^2})}{(q^2+z^2)^{(m+n+1)/2}}.$$

(In the general case, where m and n may be complex, the only restriction placed upon them is that their real parts be greater than -1). Sonine showed in addition that if we put $m=-1/2$, and make use of the relation $J_{-1/2}(t)=(2/\pi t)^{1/2}\cos t$, we obtain

$$\int_0^{\pi/2}\cos(aq\cos\phi)J_n(az\sin\phi)\sin^{n+1}\phi\, d\phi = \tfrac{1}{2}\int_0^\pi \cos(aq\cos\phi)J_n(az\sin\phi)\sin^{n+1}\phi\, d\phi = \left(\frac{\pi}{2a}\right)^{1/2} z^n\,\frac{J_{n+1/2}(a\rho)}{\rho^{n+1/2}},$$

where $\rho\equiv(q^2+z^2)^{1/2}$. Finally, if we note that the value of the last integral is zero if we replace $\cos(aq\cos\phi)$ by $\sin(aq\cos\phi)$, we have immediately an expression for (A), to wit:

$$\text{(A}'\text{)} \qquad \int_0^\pi e^{iaq\cos\phi}J_n(az\sin\phi)\sin^{n+1}\phi\, d\phi = \left(\frac{2\pi}{a}\right)^{1/2} z^n\,\frac{J_{n+1/2}(a\rho)}{\rho^{n+1/2}}.$$

It will be found that we can evaluate (B) and (C) by putting $m=1/2$ and $m=3/2$ respectively in (S), noting that $J_{1/2}(t)=(2/\pi t)^{1/2}\sin t$, $J_{3/2}(t)=(1/t)J_{1/2}(t)-J_{-1/2}(t)$. We obtain

[15] N. Sonine, *Mathematische Annalen*, vol. 16 (1880), p. 36; N. Nielsen, *Handbuch der Cylinderfunktionen* (1904), p. 181.

(B′) $\int_0^\pi e^{iaq\cos\phi} J_n(az\sin\phi)\cos\phi\sin^{n+1}\phi\, d\phi = i\left(\frac{2\pi}{a}\right)^{1/2} qz^n \frac{J_{n+3/2}(a\rho)}{\rho^{n+3/2}}$

(C′) $\int_0^\pi e^{iaq\cos\phi} J_n(az\sin\phi)\cos^2\phi\sin^{n+1}\phi\, d\phi$

$$= \left(\frac{2\pi}{a}\right)^{1/2} z^n \left[\frac{1}{a}\,\frac{J_{n+3/2}(a\rho)}{\rho^{n+3/2}} - q^2\frac{J_{n+5/2}(a\rho)}{\rho^{n+5/2}}\right].$$

To adapt these formulae to our use, we have only to put r for a, ζ for q, and $(\xi^2+\eta^2)^{1/2}$ for z. Then $\rho \equiv (q^2+z^2)^{1/2}$ becomes $(\xi^2+\eta^2+\zeta^2)^{1/2}$. We shall continue to designate this last radical by ρ, thereby conforming to our use of r for $(x^2+y^2+z^2)^{1/2}$.

The integration with respect to r presents no particular difficulty. Bessel's functions of low orders which are half of odd integers are readily reduced to trigonometric functions. Once the ten integrals are evaluated, the G_{jk} are found to be as follows:

$$G_{11} = \frac{1}{(2\pi)^{1/2}\rho^5}\Big[2(-\xi^2+\eta^2+\zeta^2)(\sin\rho\lambda - \sin\rho\mu)$$
$$-(\eta^2+\zeta^2)(\rho\lambda\cos\rho\lambda - \rho\mu\cos\rho\mu) + (2\xi^2-\eta^2-\zeta^2)\int_{\rho\mu}^{\rho\lambda}\frac{\sin u}{u}\,du\Big]$$

$G_{22} = G_{11}$ with ξ and η interchanged.
$G_{33} = G_{11}$ with ξ and ζ interchanged.

$$G_{12} = \frac{-\xi\eta}{(2\pi)^{1/2}\rho^5}\Big[4(\sin\rho\lambda - \sin\rho\mu)$$
$$-(\rho\lambda\cos\rho\lambda - \rho\mu\cos\rho\mu) - 3\int_{\rho\mu}^{\rho\lambda}\frac{\sin u}{u}\,du\Big]$$

$G_{13} = G_{12}$ with η and ζ interchanged.
$G_{23} = G_{12}$ with ξ and ζ interchanged.

$$G_{35} = \frac{-\xi}{i(2\pi)^{1/2}\rho^4}\left[2(\cos\rho\lambda - \cos\rho\mu) + (\rho\lambda\sin\rho\lambda - \rho\mu\sin\rho\mu)\right]$$

$G_{16} = G_{35}$ with ξ and η interchanged.
$G_{24} = G_{35}$ with ξ and ζ interchanged.
The matrix $\boldsymbol{G}$ is given by

$$\underset{\left(\begin{smallmatrix}\mu<\lambda<0\\ 0<\mu<\lambda\end{smallmatrix}\right)}{\boldsymbol{G}(\xi,\eta,\zeta;\Delta)} = \left\{\begin{array}{ccc|ccc} G_{11} & G_{12} & G_{13} & 0 & -G_{24} & G_{16} \\ G_{12} & G_{22} & G_{23} & G_{24} & 0 & -G_{35} \\ G_{13} & G_{23} & G_{33} & -G_{16} & G_{35} & 0 \\ \hline \multicolumn{3}{c|}{\text{upper right, signs changed}} & \multicolumn{3}{c}{\text{upper left}} \end{array}\right\}.$$

In connection with these results, we note the important fact that the matrix $\boldsymbol{G}$ is the same for the two cases $\mu < \lambda < 0$ and $0 < \mu < \lambda$. Also, we are afforded a partial check on our calculations in that since the $E'_{jk}(\Delta) \varepsilon L_2$, so also should the $\dot{G}_{jk}(\Delta)$, as functions of ξ, η, ζ, have this property. That they actually do is easy to verify.

9. *The operator* $\boldsymbol{F}(\boldsymbol{H}) \equiv e^{it\boldsymbol{H}}$. We are now prepared to show that the operator $\boldsymbol{F}(\boldsymbol{H}) \equiv e^{it\boldsymbol{H}}$ is a matrix whose components $F_{jk}(\boldsymbol{H})$ are given by the relations

$$(9.1) \qquad F_{jk}(\boldsymbol{H}) f_k(x, y, z)$$
$$= \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{+\infty} D_{jk}(\xi, \eta, \zeta; t, A) f_k(x_1, y_1, z_1)\, dx_1 dy_1 dz_1,$$
$$D_{jk}(\xi, \eta, \zeta; t, A) \equiv \left[\int_{-A}^{0-0} + \int_{0+0}^{A}\right] e^{it\lambda}\, d_\lambda G_{jk}(\xi, \eta, \zeta; \Delta),$$

wherein the G_{jk} are the elements of the matrix determined in the last section. Before setting out to prove this relation, however, we shall make a few observations regarding the calculation of the D_{jk}.

It was found in the last section that there were only nine distinct G_{jk} (with the exception of a sign), which in turn could be classified as three groups of three each, those elements in each group differing from each other only insofar as ξ, η, ζ are concerned. Hence for the moment we need consider only one element in each group—say G_{11}, G_{12}, G_{35}. We have

$$d_\lambda G_{11}(\Delta) = \frac{1}{(2\pi)^{1/2}\rho^4}\Big[(-2\xi^2 + \eta^2 + \zeta^2)\cos\rho\lambda + (\eta^2 + \zeta^2)\rho\lambda\sin\rho\lambda$$
$$+ (2\xi^2 - \eta^2 - \zeta^2)\frac{\sin\rho\lambda}{\rho\lambda}\Big]\, d\lambda$$
$$d_\lambda G_{12}(\Delta) = \frac{-\xi\eta}{(2\pi)^{1/2}\rho^4}\left[3\cos\rho\lambda + \rho\lambda\sin\rho\lambda - 3\frac{\sin\rho\lambda}{\rho\lambda}\right] d\lambda$$
$$d_\lambda G_{35}(\Delta) = \frac{-\xi}{i(2\pi)^{1/2}\rho^3}\,[\rho\lambda\cos\rho\lambda - \sin\rho\lambda]\, d\lambda.$$

Now all the terms involved above, multiplied by $e^{it\lambda}$, are continuous at the origin. Hence, we may replace $\left[\int_{-A}^{0-0} + \int_{0+0}^{A}\right]$ by $\int_{-A}^{A}$. The calculation of the D_{jk} is a relatively simple matter. We shall not list the results here, as it seems advisable, for all purposes of reference, to include them in the summary (Section 11). It is a matter of simple routine to show that, as functions of x_1, y_1, z_1, the $D_{jk} \varepsilon L_2$.

We have now to verify the relation (9. 1). Among the important characteristic properties of a resolution of the identity are the following: [16]

$$\underset{\lambda\to+\infty}{\text{l.i.m.}}\, \boldsymbol{E}(\lambda)f = f, \qquad \underset{\lambda\to-\infty}{\text{l.i.m.}}\, \boldsymbol{E}(\lambda)f = 0$$

for any function f in its domain. In the case at hand, then,

$$\underset{A\to\infty}{\text{l.i.m.}}\, [\boldsymbol{E}(A) - \boldsymbol{E}(-A)]f = f.$$

Now by Stone, Theorem 6. 6, $e^{it\cdot}$ is a unitary transformation.[17] Hence we must have

$$e^{itH}f = \underset{A\to\infty}{\text{l.i.m.}}\, e^{itH}[\boldsymbol{E}(A) - \boldsymbol{E}(-A)]f, \tag{9. 2}$$

for every $f \,\varepsilon\, L_{2,6}$. Consider now

$$[\boldsymbol{E}(A) - \boldsymbol{E}(-A)]f = \{[\boldsymbol{E}(A) - \boldsymbol{E}(0+0)] - [\boldsymbol{E}(-A) - \boldsymbol{E}(0-0)]\}f \\ + [\boldsymbol{E}(0+0) - \boldsymbol{E}(0-0)]f, \qquad A > 0.$$

Regarding the last term on the right-hand side of this relation, it was pointed out in Section 7 that the range of the operator $[\boldsymbol{E}(0+0) - \boldsymbol{E}(0-0)]$ consists of those and only those functions which are characteristic elements of $\boldsymbol{H}$ corresponding to the characteristic value $l = 0$. The term in question can thus be eliminated from our present consideration. For if f is a characteristic element, then $\boldsymbol{H}f = 0$. Referring back to our original problem—that of determining the solution of the equation $\partial f/\partial t = i\boldsymbol{H}f$—it is evident that in such a case $\partial f/\partial t = 0$, and the function f is constant in time. Such functions we can afford to disregard in the present analysis.[18] It should be pointed out, however, that we do not eliminate the term $[\boldsymbol{E}(0+0) - \boldsymbol{E}(0-0)]f$ as a matter of mere convenience, but because of the serious difficulties it would cause in the work which is to follow. We shall hence forth understand, when we write $\boldsymbol{E}(A)$ and $\boldsymbol{E}(-A)$, that we mean $[\boldsymbol{E}(A) - \boldsymbol{E}(0+0)]$ and $[\boldsymbol{E}(-A) - \boldsymbol{E}(0-0)]$ respectively.

Now let $\phi_n(t, \lambda)$ be a function which for t fixed converges uniformly to $e^{it\lambda}$ on $(-A, A)$. Then in accordance with Stone, Theorem 6. 1,

$$e^{itH}[\boldsymbol{E}(A) - \boldsymbol{E}(-A)]f = \underset{n\to\infty}{\text{l.i.m.}}\, \phi_n(t, \lambda)[\boldsymbol{E}(A) - \boldsymbol{E}(-A)]f. \tag{9. 3}$$

---

[16] Stone, Definition 5. 1.

[17] The reader should note the revision of notation brought about in Stone, Definition 6. 4. In accordance with this revision, we are using $F(H)$ in the place of $T(F)$ in Theorem 6. 1.

[18] See also Section 10.

We choose $\phi_n(t,\lambda)$ as the step-function defined as follows. Divide the interval $(-A, A)$ into n parts, the only restriction being that the origin be interior to no sub-interval. This is, of course, no real restriction, since we could just as well work with the two intervals $(-A, 0-0)$ and $(0+0, A)$ separately. Put

$$\phi_n(t,\lambda) = \begin{cases} e^{i\lambda_\nu t}, & \lambda_\nu \leqq \lambda \leqq \lambda_{\nu+1} \qquad (\nu = 1, \cdots, n), \\ 0, & \text{elsewhere.} \end{cases}$$

It is apparent that $\phi_n(t,\lambda)$ has the property which we required of it above. Put for simplicity

$$W(H; A) \equiv e^{itH}[E(A) - E(-A)],$$

and denote as usual its elements by $W_{jk}(H; A)$. Since the relation (9.3) holds for an arbitrary $f \varepsilon L_{2,6}$, we must have, for the separate components,

$$W_{jk}(H; A) f_k = \underset{n\to\infty}{\text{l.i.m.}} \left\{ \sum_{\nu=0}^{n} e^{i\lambda_\nu t}[E_{jk}(\lambda_{\nu+1}) - E_{jk}(\lambda_\nu)] \right\} f_k.$$

But $E_{jk}(\lambda_{\nu+1}) - E_{jk}(\lambda_\nu)$ is nothing other than $E_{jk}(\Delta)$ where $\lambda = \lambda_{\nu+1}$, $\mu = \lambda_\nu$. If we denote this by $E_{jk}(\Delta_\nu)$, we have in view of (9.3),

$$\text{(9.4)} \quad W_{jk}(H; A) f_k = \underset{n\to\infty}{\text{l.i.m.}} \frac{1}{(2\pi)^{3/2}} \int\int_{-\infty}^{+\infty}\int \sum_{\nu=0}^{n} {}^{i\lambda_\nu t} G_{jk}(\xi,\eta,\zeta;\Delta_\nu) \cdot f_k(x_1,y_1,z_1)\, dx_1 dy_1 dz_1$$

$$= \underset{n\to\infty}{\text{l.i.m.}} \frac{1}{(2\pi)^{3/2}} \int\int_{-\infty}^{+\infty}\int \left\{ \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] \phi_n(t,\lambda) \cdot d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right\} f_k(x_1,y_1,z_1)\, dx_1 dy_1 dz_1.$$

It remains to show that (9.4) implies

$$\text{(9.5)} \quad W_{jk}(H; A) f_k = \frac{1}{(2\pi)^{3/2}} \int\int_{-\infty}^{+\infty}\int \left\{ \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] e^{it\lambda} \cdot d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right\} f_k(x_1,y_1,z_1)\, dx_1 dy_1 dz_1$$

almost everywhere.

We introduce the functions f_k^B which are identically equal to f_k inside and on an arbitrary origin-centered, axis-parallel cube of side $2B$, and zero elsewhere. We then consider the integral

$$(9.6)\qquad \iiint_{-\infty}^{+\infty} \left\{ \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] [e^{it\lambda} - \phi_n(t,\lambda)] d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right\} \cdot f_k^B(x_1,y_1,z_1)\, dx_1 dy_1 dz_1.$$

Since for t fixed, $\phi_n(t,\lambda)$ converges uniformly to $e^{it\lambda}$ on $(-A, A)$, there exists a sequence of constants $\{M_n\}$ depending only on n, such that (i) $|e^{it\lambda} - \phi_n(t,\lambda)| \leq M_n$ for every n, and (ii) $\lim_{n\to\infty} M_n = 0$. Hence

$$\left| \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] [e^{it\lambda} - \phi_n(t,\lambda)] d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right| \leq M_n V_{-A}^{A}(G_{jk}) \leq M_n \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] |G'_{jk}(\xi,\eta,\zeta;\Delta)|\, d\lambda,$$

where the prime denotes differentiation with respect to λ. It is easy to show that the integral on the right-hand side of this relation exists as a function of x_1, y_1, z_1, and is of integrable square over the cube of side $2B$. If we denote the integral in question by $N_{jk}(\xi,\eta,\zeta;A)$, then for fixed x, y, z, the integral (9.6) exists and is dominated by

$$M_n \iiint_{-\infty}^{+\infty} N_{jk}(\xi,\eta,\zeta;A)\, |f_k^B(x_1,y_1,z_1)|\, dx_1 dy_1 dz_1.$$

This observation, together with the fact that $M_n \to 0$ as $n \to \infty$ insures the convergence of (9.6) to zero as $n \to \infty$. In that case,

$$(9.7)\qquad \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{+\infty} \left\{ \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] e^{it\lambda} \cdot d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right\} f_k^B(x_1,y_1,z_1)\, dx_1 dy_1 dz_1$$
$$= \lim_{n\to\infty} \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{+\infty} \left\{ \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] \phi_n(t,\lambda) \cdot d_\lambda G_{jk}(\xi,\eta,\zeta;\Delta) \right\} f_k^B(x_1,y_1,z_1)\, dx_1 dy_1 dz_1$$

for all x, y, z. Now the terms under l.i.m. and lim in (9.4) and (9.7) respectively are identical. Hence, by a well known theorem regarding the relation between limits in the mean and ordinary limits, the left-hand sides of these equations must be equal almost everywhere. That is, the relation (9.5) is true for all functions f_k^B. Let us write this in the form

$$(9.8)\quad W_{jk}(\boldsymbol{H}; A) f_k^B = \frac{1}{(2\pi)^{3/2}} \int\int\int_{-\infty}^{+\infty} D_{jk}(\xi, \eta, \zeta; t, A) f_k^B(x_1, y_1, z_1)\, dx_1 dy_1 dz_1,$$

$$D_{jk}(\xi, \eta, \zeta; t, A) \equiv \left[\int_{-A}^{0-0} + \int_{0+0}^{A} \right] e^{it\lambda} d_\lambda G_{jk}(\xi, \eta, \zeta; \Delta),$$

almost everywhere.

We now remove the restrictions on f_k. To do so, we note that $f_k = \underset{B\to\infty}{\text{l.i.m.}}\, f_k^B$. It is then not difficult to show that $W_{jk}(\boldsymbol{H}; A) f_k = \underset{B\to\infty}{\text{l.i.m.}}\, W_{jk}(\boldsymbol{H}; A) f_k^B$. In addition, since the $D_{jk} \varepsilon L_2$ as functions of x_1, y_1, z_1, we must have, for fixed x, y, z,

$$\frac{1}{(2\pi)^{3/2}} \int\int\int_{-\infty}^{+\infty} D_{jk} f_k dx_1 dy_1 dz_1 = \lim_{B\to\infty} \int\int\int_{-\infty}^{+\infty} D_{jk} f_k^B dx_1 dy_1 dz_1.$$

By the same argument just used regarding l.i.m. and lim, the relation (9.8) must hold for all functions $f_k \varepsilon L_2$. This, together with (9.2), completes the proof of (9.1).

10. *The equations* $\nabla \cdot \boldsymbol{e} = 0$ *and* $\nabla \cdot \boldsymbol{h} = 0$. Recall that of the four electrodynamic field equations, we have so far failed to consider the two equations $\nabla \cdot \boldsymbol{e} = 0$ and $\nabla \cdot \boldsymbol{h} = 0$. We propose in this section to investigate their significance.

We discussed in Section 7 a linear manifold dense in P, the characteristic manifold of $\boldsymbol{H}$. This was the set P_0, consisting of all functions $\boldsymbol{f}$ of the form

$$\boldsymbol{f} = \left(\frac{\partial p_1}{\partial x}, \frac{\partial p_1}{\partial y}, \frac{\partial p_1}{\partial z}; \frac{\partial p_2}{\partial x}, \frac{\partial p_2}{\partial y}, \frac{\partial p_2}{\partial z} \right),$$

where p_1 and p_2 belong to L_2, are absolutely continuous in x, y, and z separately, and have the property that their first partial derivatives with respect to x, y, or z belong to L_2. Let us examine the relation between the orthogonal complements of P and P_0. If we denote these manifolds by Q and Q_0 respectively, it is apparent that, since $P_0 \subset P$, $Q_0 \supseteq Q$. To show that $Q_0 = Q$, we argue as follows:

Since any function in $L_{2,6}$ can be uniquely represented as the sum of two functions, one in P and the other in Q, we have only to show that no function of $P - P_0$ is orthogonal to every function of P_0. Suppose, then, that $\boldsymbol{f} \varepsilon P - P_0$. Then a sequence of functions $\{\boldsymbol{f}_n\} \varepsilon P_0$ can be found such that $|\boldsymbol{f}_n - \boldsymbol{f}|^2 \equiv |\boldsymbol{f}_n|^2 + |\boldsymbol{f}|^2 - (\boldsymbol{f}, \boldsymbol{f}_n) - (\boldsymbol{f}_n, \boldsymbol{f}) \to 0$ as $n \to \infty$. Now assume $\boldsymbol{f}$ orthogonal to every function of P_0. Then the last relation becomes

$|f_n|^2 + |f|^2 \to 0$ as $n \to \infty$. This can be true only for $f \equiv 0$. But the null element belongs to Q as well as to P. Hence the orthogonal complement of P_0 coincides with that of P.

The set $Q_0 = Q$ must consist of all functions $g \varepsilon L_{2,6}$ such that

$$\int\int\int_{-\infty}^{+\infty} g \cdot \bar{f}\, dxdydz = \lim_{A\to\infty} \int\int\int_{-A}^{A} g \cdot \bar{f}\, dxdydz = 0 \tag{10.1}$$

for every $f \varepsilon P_0$. If the second term of this relation be integrated by parts, we obtain

$$\begin{aligned}\lim_{A\to\infty} \Bigg\{ &\int\int_{-A}^{A} [g_1\bar{p}_1]_{x=-A}^{x=A} dydz + \int\int_{-A}^{A} [g_2\bar{p}_1]_{y=-A}^{y=A} dxdz + \int\int_{-A}^{A} [g_3\bar{p}_1]_{z=-A}^{z=A} dxdy \\ &+ 3 \text{ similar terms involving the products of } \bar{p}_2 \text{ and } g_4, g_5, g_6, \text{ respectively} \\ &- \int\int\int_{-A}^{A} \Big[\Big(\frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z} \Big) \bar{p}_1 + \Big(\frac{\partial g_4}{\partial x} + \frac{\partial g_5}{\partial y} + \frac{\partial g_6}{\partial z} \Big) \bar{p}_2 \Big] dxdydz \Bigg\} = 0.\end{aligned} \tag{10.2}$$

It is now possible to pick out a sequence $\{a\}$ of A's such that the first six terms of this expression vanish in the limit. To show this, we argue as follows.

Consider any one of the terms in question—say the first. Since g_1 and $\bar{p}_1$ both belong to L_2, their product, which we shall denote by q, belongs to L_1. Then for a given ϵ, we can find a b so large that

$$\int_b^\infty \Big[\int\int_{-\infty}^{+\infty} |q(x, y, z)|\, dydz \Big] dx < \epsilon^2;$$

in which case, a fortiori,

$$\int_b^\infty \phi(x, c)\, dx \equiv \int_b^\infty \Big[\int\int_{-c}^{c} |q(x, y, z)|\, dydz \Big] dx < \epsilon^2.$$

It follows that the set of values of x, $x > b$, for which $\phi(x, c) \geqq \epsilon$ is of measure $< \epsilon$. That is, outside a set of values of x, $b < x \leq \infty$, of measure $< \epsilon$, $\phi(x, c) < \epsilon$. If now we take $c > b$, we can choose an a, $b < a < c$, so that $\phi(a, c) < \epsilon$. It follows that

$$\Big| \int\int_{-a}^{a} q(a, y, z)\, dydz \Big| \leq \int\int_{-a}^{a} |q(a, y, z)|\, dydz < \epsilon,$$

and as $\epsilon \to 0$ over any sequence of values, we can find corresponding values of $a \to \infty$. Now the first term of (10. 2) is of the form

$$\lim_{A\to\infty} \left\{ \int\int_{-A}^{A} q(A, y, z)\, dydz - \int\int_{-A}^{A} q(-A, y, z)\, dydz \right\}.$$

It is evident not only that the sequence $\{a\}$ of A's can be so chosen that both parts of this first term vanish simultaneously in the limit, but that it can be so chosen that the same will be true of both parts of each of the first six terms of (10. 2).

The condition (10. 1) is consequently equivalent to

$$\int\int\int_{-\infty}^{+\infty} \left[\left(\frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z} \right) \bar{p}_1 + \left(\frac{\partial g_4}{\partial x} + \frac{\partial g_5}{\partial y} + \frac{\partial g_6}{\partial z} \right) \bar{p}_2 \right] dxdydz = 0.$$

This in turn is equivalent to

$$\frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z} = 0, \qquad \frac{\partial g_4}{\partial x} + \frac{\partial g_5}{\partial y} + \frac{\partial g_6}{\partial z} = 0$$

almost everywhere; or if $\boldsymbol{g} \equiv (e_x, e_y, e_z, h_x, h_y, h_z)$, equivalent to the conditions $\nabla \cdot \boldsymbol{e} = 0$, $\nabla \cdot \boldsymbol{h} = 0$ almost everywhere.

We are thus led to the conclusion that of all functions belonging to $L_{2,6}$, those and only those which satisfy the divergence conditions constitute the orthogonal complement of the characteristic manifold of $\boldsymbol{H}$.

11. *Summary.* The object of this paper has been to obtain rigorously a general integral representation for the solution of the classical electrodynamic field equations in the case of empty space containing no charges or currents. These equations were expressed in the form $\nabla \cdot \boldsymbol{e} = 0$, $\nabla \cdot \boldsymbol{h} = 0$, $\nabla \times \boldsymbol{e} = -\partial \boldsymbol{h}/\partial t$, $\nabla \times \boldsymbol{h} = \partial \boldsymbol{e}/\partial t$. The significance of the two divergence equations was discussed in the last section. It was found that upon putting $\boldsymbol{v} \equiv (e_x, e_y, e_z, e_x, h_y, h_z)$, the two curl equations could be written as the single matrix equation

(11. 1) $$\partial \boldsymbol{v}/\partial t = i\boldsymbol{H}\boldsymbol{v},$$

$$
\boldsymbol{H} \equiv \left\{ \begin{array}{c|c} \mathbf{0} & \begin{array}{ccc} 0 & -\frac{i}{1}\frac{\partial z}{\partial} & \frac{i}{1}\frac{\partial y}{\partial} \\ \frac{1}{i}\frac{\partial}{\partial z} & 0 & -\frac{1}{i}\frac{\partial}{\partial x} \\ -\frac{1}{i}\frac{\partial}{\partial y} & \frac{1}{i}\frac{\partial}{\partial x} & 0 \end{array} \\ \hline \begin{array}{c}\text{upper right,}\\ \text{signs changed}\end{array} & \mathbf{0} \end{array} \right\}.
$$

The domain of the operator $\boldsymbol{H}$ was not precisely determined, but it is sufficient for all practical purposes to know that it contains the class of all vector functions $\boldsymbol{v}$ whose components along with their first partial derivatives with respect to x, y, or z belong to L_2 over all space, and whose components in addition are absolutely continuous in x, y, and z separately.

The solution of the equation (11. 1) was found to be

$$
\begin{aligned}
\boldsymbol{v} &= e^{itH}\boldsymbol{v}_0 \\
&= \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{(2\pi)^{3/2}} \int\int_{-A}^{A}\int \boldsymbol{D}(\xi, \eta, \zeta; t, A)\, \boldsymbol{v}_0(x_1, y_1, z_1)\, dx_1 dy_1 dz_1,
\end{aligned}
$$

where $\xi \equiv x_1 - x$, $\eta \equiv y_1 - y$, $\zeta \equiv z_1 - z$, and where

$$
\boldsymbol{D}(\xi, \eta, \zeta; t, A) \equiv \left\{ \begin{array}{c|c} \begin{array}{ccc} D_{11} & D_{12} & D_{13} \\ D_{12} & D_{22} & D_{23} \\ D_{13} & D_{23} & D_{33} \end{array} & \begin{array}{ccc} 0 & -D_{24} & D_{16} \\ D_{24} & 0 & -D_{35} \\ -D_{16} & D_{35} & 0 \end{array} \\ \hline \begin{array}{c}\text{upper right,}\\ \text{signs changed}\end{array} & \begin{array}{c}\text{upper}\\ \text{left}\end{array} \end{array} \right\}.
$$

The elements of this matrix are given by the following expressions, wherein $\rho \equiv (\xi^2 + \eta^2 + \zeta^2)^{1/2}$.

$$
\begin{aligned}
D_{11} = \frac{1}{(2\pi)^{1/2}\rho^4} \Bigg\{ & \frac{t(-2\xi^2 + \eta^2 + \zeta^2) + 2\rho(-\xi^2 + \eta^2 + \zeta^2)}{(t+\rho)^2} \sin (t+\rho)A \\
& + \frac{t(-2\xi^2 + \eta^2 + \zeta^2) - 2\rho(-\xi^2 + \eta^2 + \zeta^2)}{(t-\rho)^2} \sin (t-\rho)A \\
& - A\rho(\eta^2 + \zeta^2) \left[\frac{\cos(t+\rho)A}{t+\rho} - \frac{\cos(t-\rho)A}{t-\rho} \right] \\
& + \frac{2\xi^2 - \eta^2 - \zeta^2}{\rho} \int_{-A}^{A} e^{it\lambda} \frac{\sin \rho\lambda}{\lambda}\, d\lambda \Bigg\}.
\end{aligned}
$$

$D_{22} = D_{11}$ with ξ and η interchanged.
$D_{33} = D_{11}$ with ξ and ζ interchanged.

$$D_{12} = \frac{-\xi\eta}{(2\pi)^{1/2}\rho^4}\left\{\frac{3t+4\rho}{(t+\rho)^2}\sin(t+\rho)A + \frac{3t-4\rho}{(t-\rho)^2}\sin(t-\right.$$
$$\left. -A\rho\left[\frac{\cos(t+\rho)A}{t+\rho} - \frac{\cos(t-\rho)A}{t-\rho}\right] - \frac{3}{\rho}\int_{-A}^{A} e^{it\lambda}\right.$$

$D_{13} = D_{12}$ with η and ζ interchanged.
$D_{23} = D_{12}$ with ξ and ζ interchanged.

$$D_{35} = \frac{-\xi}{(2\pi)^{1/2}\rho^3}\left\{\frac{t+2\rho}{(t+\rho)^2}\sin(t+\rho)A - \frac{t-2\rho}{(t-\rho)^2}\sin(t-\right.$$
$$\left. -A\rho\left[\frac{\cos(t+\rho)A}{t+\rho} + \frac{\cos(t-\rho)A}{t-\rho}\right]\right\}$$

$D_{16} = D_{35}$ with ξ and η interchanged.
$D_{24} = D_{35}$ with ξ and ζ interchanged.

Finally, if the results are desired in Heaviside-Lorentz ι to replace t by ct, where c is the velocity of light *in vacuo.*

YALE UNIVERSITY.

CURVATURE TRAJECTORIES.[1]

By George Comenetz.

1. The geometrical objects which we study in this paper are triply infinite families of plane curves. We deal with three interesting special types of such families: the *dynamical, sectional,* and *curvature* types. These are related, in that all of their families possess the differential property which Kasner has given as dynamical property I (quoted below). Moreover, the body of families belonging to each one of the types is invariant under the group of all projective point transformations. A first set of three problems is suggested by the common differential property: to determine what families of curves appear under some two of the three types at once. A second set of three problems may be derived from the projective character: to determine which families under each type are entirely composed of conic sections. For two questions in each set the solutions are already known. The remaining pair of questions, connected with the curvature type, are answered below. A remarkable feature of this group of problems is that in every case an explicit answer can be obtained.

I wish to thank Professor Kasner for setting these problems, and for his assistance in solving them.

2. First we define the three types of families in question.

With an arbitrarily given positional field of force there is associated a family of *dynamical trajectories.* These are the totality of curves along which a mass particle can move under the influence of the field. Since the particle can start from any point, in any direction, and with any velocity, it can traverse ∞^3 different paths. The differential equation of these ∞^3 curves is

$$(1) \qquad y''' = \frac{\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2}{\psi - y'\phi} y'' - \frac{3\phi}{\psi - y'\phi} y''^2,$$

where $\phi(x, y)$ and $\psi(x, y)$ are the components of force.[2]

The most familiar example of a dynamical family is the set of all conics with a common focus, the orbits in the Newtonian field.

[1] Abstract in *Bulletin of the American Mathematical Society,* March 1934, p. 213.

[2] E. Kasner, "The trajectories of dynamics," *Transactions of the American Mathematical Society,* vol. 7 (1906), pp. 401-424. Also "Differential-geometric aspects of dynamics," *Princeton Colloquium Lectures on Mathematics* (1913; new edition 1934).

The projective character of dynamical families was discovered by Appell.[3]

A *sectional* family is constructed by taking an arbitrary surface in space, drawing all the plane curves possible on the surface, and projecting these curves from any fixed center onto the (x, y)-plane. A triply infinite family of curves is thus produced, since there are ∞^3 plane sections of the surface. The projective nature of the construction is obvious. If the center of projection is taken to be the point at infinity on the z-axis, the projection becomes orthogonal, and the equation of the family of projected curves is found to be

$$(2)\quad y''' = \frac{f_{xxx} + 3f_{xxy}y' + 3f_{xyy}y'^2 + f_{yyy}y'^3}{f_{xx} + 2f_{xy}y' + f_{yy}y'^2}\, y'' + \frac{3(f_{xy} + f_{yy}y')}{f_{xx} + 2f_{xy}y' + f_{yy}y'^2}\, y''^2,$$

where $z = f(x, y)$ represents the surface.[4]

For instance, stereographic projection of the sphere produces the family of all circles.

Curvature trajectories may be described as an analogue for three-parameter families of what isogonal trajectories are for two-parameter families. Here we take as a basis an arbitrarily given *doubly* infinite family of curves. A new curve drawn at random would be tangent at each of its points to some curve of the base family. If now we require the new curve, say Γ, to be drawn in such a way that the ratio of its curvature at any point, to the curvature of the member of the base family which it touches at that point, shall be kept a fixed number c as we move along Γ, then we have a curvature trajectory of the base family. For any one value of c, ∞^2 trajectories can be drawn; hence ∞^3 in all. That the construction is projective follows from Mehmke's theorem, that the ratio of curvatures of curves tangent at a point is a projective invariant. The differential equation of a curvature family has the form

$$(3)\qquad y''' = [(F_x + F_y y')/F]y'' + (F_{y'}/F)y''^2,$$

where $y'' = F(x, y, y')$ represents the arbitrary base family.[5]

For instance, if the base family is composed of all the unit circles, the curvature trajectories will be the family of all circles.

---

[3] "De l'homographie en mecanique," *American Journal of Mathematics*, vol. 12 (1890), p. 103.

[4] Kasner, "Dynamical trajectories and the ∞^3 plane sections of a surface," *Proceedings of the National Academy of Sciences*, vol. 17 (1931), p. 370.

[5] Kasner, "Dynamical trajectories and curvature trajectories," *Bulletin of the American Mathematical Society*, vol. 40 (1934), p. 449. The ∞^2 straight lines of the plane appear as the one degenerate family under all three types. This family is excluded in the following.

We observe that for each of the three types just defined the differential equation has the following special form in y'':

$$y''' = G(x, y, y')y'' + H(x, y, y')y''^2. \tag{4}$$

This special form is the analytic expression of the first of Kasner's set of six geometrical properties for dynamical families:

Property I. If to each of the ∞^1 curves having a given lineal element in common the osculating parabola is drawn at that element, the foci will lie on a circle through the point of the element.

The set of all families obeying Property I is once more a projective body. In fact equation (4) is invariant under the full projective group, correlations as well as collineations.[6]

3. We can now summarize the answers to the two sets of problems.

A1). The families of curves which are both dynamical and sectional are those derived from *cones* (of any cross section). They are the trajectories of a class of central fields of force of inverse square character.[7]

A2) The families of curves which are at once of the dynamical and curvature types are the trajectories of exactly all *central fields of force.*

A3) The families of curves which are of both the sectional and curvature types are those derived from *cones* and *quadric surfaces.*

B1) There are three cases in which all the trajectories in a field of force are conic sections:

a) Conics with a common polar pair (the point not on the line).
b) Conics tangent to two intersecting lines.
c) Conics tangent to a line at a point.

B2) There are four cases in which all the curves of a sectional family are conic sections. These are b) and c), and:

d) Conics through two fixed points.
e) Conics tangent twice to a fixed (proper) conic.

B3) There are five cases in which all the curves of a curvature family are conics. These are a), b), c), d), and e).

Equation (1) for dynamical trajectories, the definition of the sectional type, and the concept of curvature trajectories, as well as the solutions of problems A1, A2, and B2 are all due to Kasner. B1 is the well-known

[6] *Princeton Colloquium*, p. 77.

[7] The force equals $f(\theta)/r^2$. Without recognizing the connection with surfaces, Jacobi discussed these fields, showing that their trajectories are obtainable by quadratures. (*Werke*, IV, p. 282.)

Bertrand's problem, solved by Darboux and Halphen.[8] A3 and B3 are proved below, and we also discuss B2.

4. *Proof of* A3. We remark first that part of A3 is implied at once by the combination of A1 and A2, and that another part may be deduced from the circumstance that the family of all circles is an example of a sectional family and also of a curvature family.

If a family of curves belongs to both the sectional and curvature types, equations (2) and (3) must become identical.[9] We may therefore equate the coefficients of y'' and of y''^2. Integrating the latter equation, we find

$$F(x, y, y') = (f_{xx} + 2f_{xy}y' + f_{yy}y'^2)^{3/2}\,\theta(x, y), \tag{5}$$

where $\theta(x, y)$ is arbitrary. We use this relation to eliminate F from the result of equating the coefficients of y''. After simplifying, we obtain a cubic polynomial in y' which must vanish identically. The four coefficients therefore vanish separately, and we have a system of four equations in f and θ, two of which are

$$\bar{\theta}_x = -\tfrac{1}{2}f_{xxx}/f_{xx}, \qquad \bar{\theta}_y = -\tfrac{1}{2}f_{yyy}/f_{yy}, \tag{6}$$

where $\bar{\theta} = \log\theta$. Using these two, we eliminate θ, and thus find that the following conditions must be satisfied by f:

$$2f_{xy}f_{yy}f_{xxx} - 3f_{xx}f_{yy}f_{xxy} + f^2_{xx}f_{yyy} = 0, \tag{7_1}$$

$$f^2_{yy}f_{xxx} - 3f_{xx}f_{yy}f_{xyy} + 2f_{xx}f_{xy}f_{yyy} = 0, \tag{7_2}$$

$$(f_{xxx}/f_{xx})_y = (f_{yyy}/f_{yy})_x. \tag{7_3}$$

The details of the elimination show that equations (7) are sufficient as well as necessary for the identity of (2) and (3). Hence the surfaces $z = f(x, y)$ which afford solutions of our problem are those defined by the system (7).

By differentiating (7_1) and (7_2) partially with respect to x and y, and combining, we find that they imply (7_3), unless

$$f_{xx}f_{yy} - f^2_{xy} = 0; \tag{8}$$

that is, unless the surface is developable. On the other hand, (7_1) and (7_2) are themselves found to be consequences of (8). Thus there are two cases:

---

[8] *Comptes Rendus*, vol. 84 (1877), pp. 671, 731, 760, 936, 939. References in *Enc. der Math. Wiss.*, vol. IV 6, p. 498.

[9] A family derived by central projection from a surface S can always be considered as derived orthogonally from a surface S' projectively related to S.

We shall assume that the coefficients of y'' and y''^2 in equations (1), (2), (3), and (4) are analytic in x, y, and y'.

either the surface $z = f(x, y)$ is developable, and obeys (7_3) and (8); or it is not developable, and obeys (7_1) and (7_2).

In the first case we merely verify that (7_3) is equivalent under (8) to the condition $(f_{xxy}/f_{xy})_y = (f_{yyy}/f_{yy})_x$, which was obtained with (8) in the solution of problem A1. Hence the surfaces defined by (7_3) and (8) are those which appear in the answer to A1; namely, cones (and cylinders).

Equations (7_1) and (7_2) were derived by Hermite [10] and further discussed by Halphen.[11] Their solutions are quadrics and developables. Hence the solutions which they contribute to our problem are just quadrics. The conclusion is that the surfaces which generate curvature trajectories are cones and quadrics, as A3 states.

Proofs of A3 by direct integration of the differential equations have been given by M. Halperin and by the writer.

5. *Problem* B2. If every plane section of a surface X projects into a conic, every plane section of X is itself a conic. Then X must be a quadric surface. This may be shown, for instance, in the following way.

Let C_1 and C_2 be two conics on X, intersecting in two points J and K, and let P be a further point on X. A quadric Q can be passed through the configuration C_1C_2P. (We can take J, K, P, two other points M and N on the conics, and the tangent planes at J and K as determining elements.) Every plane through P which meets C_1 and C_2 in four points must cut out the same conic on Q as on X, since five points determine a conic. It follows that X must be Q.

The quadric may be proper or degenerate, and the center of projection may be on the surface or off. This accounts for the four projectively distinct cases b), c), d), and e); (it is easy to show synthetically that case e), for example, is obtained by projecting a proper quadric from an outside point). A unified description of these would be: conics twice tangent to a fixed curve of second order or second class, proper or degenerate; (the two most degenerate types may be said to appear coincidently in c)). The fixed locus, which may be imaginary, comes from the intersection of the quadric with the polar plane of the center of projection. It is the umbral curve, or boundary of the geometrical shadow, of the surface.

6. *Proof of* B3. In view of the previous results, a) to e) are already

---

[10] *Cours d'Analyse*, vol. 1, p. 149. Hermite notes that (7_1) and (7_2) are the conditions for the denominator of the coefficient of y'' in (2) to divide the numerator.

[11] "Sur le contact des surfaces," *Bulletin de la Société Mathématique de France*, vol. 3 (1874), p. 28. Developables enter when the polynomial B (p. 33) is a perfect square.

known to be curvature trajectories, (for a), b) and c) in B1 do come from central fields of force). The effect of this proof is, then, to show that there are no other families of conics which are curvature trajectories.

We employ the differential equation of all conics:

$$9y''^2y^{v} - 45y''y'''y^{iv} + 40y'''^3 = 0. \tag{9}$$

If an equation of the form $y''' = Gy'' + Hy''^2$ represents a family of conics, it must render (9) an identity in x, y, y', y''. We therefore differentiate to find y^{iv} and y^{v}, and substitute in (9). The result is cubic in y''.[12] Setting the four coefficients equal to zero, we have:

$$H_{y'y'} + 2HH_{y'} + \tfrac{4}{9}H^3 = 0, \tag{10_1}$$

$$G_{y'y'} + GH_{y'} + \tfrac{1}{3}GH^2 + 2H_{xy'} + 2H_{yy'}y' + HH_x + HH_yy' + H_y = 0, \tag{10_2}$$

$$H_{xx} + 2H_{xy}y' + H_{yy}y'^2 - G_xH - G_yHy' + \tfrac{1}{3}G^2H + 2G_{xy'} + 2G_{yy'}y' - GG_{y'} + G_y = 0, \tag{10_3}$$

$$G_{xx} + 2G_{xy}y' + G_{yy}y'^2 - 2GG_x - 2GG_yy' + \tfrac{4}{9}G^3 = 0. \tag{10_4}$$

A cubic which vanishes for four values of the variable vanishes identically. Consequently:

If $4\infty^2$ out of the ∞^3 curves of a family having Property I are conics, the rest are conics also.[13]

Now in (10_1) only derivatives with respect to y' appear. Hence we can solve (10_1) as an ordinary differential equation, to find how H involves y'. The result is

$$H = 3(\sigma + \tau y')/(\rho + 2\sigma y' + \tau y'^2), \tag{11_1}$$

where ρ, σ, τ are arbitrary functions of x and y. If $\rho\tau - \sigma^2 = 0$, the numerator divides the denominator, and (11_1) can be reduced to

$$H = 3/(y' - \omega), \tag{11_2}$$

where $\omega(x, y)$ is arbitrary. (The case $H = 0$ is brought under (11_2) with $\omega = 0$ by interchanging x and y.)

[12] A factor y''^3 is dropped. (9) may be looked on as the condition for sextactic points. When (10_1) holds, one root of the cubic becomes infinite; then when (10_2) holds, two roots are infinite; etc. ($y'' = \infty$ represents a *point union.*)

[13] The ∞^2 straight lines, present in every Property I family, are not to be counted in the "$4\infty^2$". This can be strengthened to "$3\infty^2$" for the dynamical type or the sectional type, since both of these satisfy (10_1) identically. J. N. Hazzidakis gave the following theorem for the dynamical type: if for two directions of initial motion all the trajectories are conics, then the family is entirely composed of conics. (*Journal für Mathematik,* vol. 133 (1908), p. 68.)

These formulas have geometric meaning. The second one expresses Property II of Kasner's set for dynamical trajectories:

Property II. There exists for each point (x, y) of the plane a certain direction ω (that of the "*force*") such that the angle between this direction and the tangent to the focal circle (of Property I) corresponding to any element (x, y, y') at the given point, is bisected by that element.

It can be shown that formula (11_1) generalizes Property II in the following way:

There exist for each point (x, y) of the plane two directions ω_1 and ω_2 (the "*asymptotic*" directions) such that the angle between the tangent at (x, y) to the focal circle corresponding to any element (x, y, y'), and the harmonic conjugate of y' with respect to ω_1 and ω_2, is bisected by the element.[14]

When ω_1 and ω_2 coincide, this reduces to Property II. Evidently (11_1) holds for the sectional type (2). The integral curves of the directions ω_1, ω_2 are then the projections of the asymptotic lines on the surface (an *asymptotic net*), and the two harmonic conjugates come from conjugate directions on the surface.

The two forms (11) for H divide the problem naturally into what we may call a *sectional case* and a *dynamical case.*

Dynamical case. We substitute (11_2) into the second of equations (10), and obtain an ordinary differential equation for G as a function of y'. Solving, we find that

$$G = (\lambda y'^2 + \mu y' + \nu)/(y' - \omega), \tag{12}$$

where λ, μ, ν, ω are any functions of x and y satisfying

$$\lambda\omega^2 + \mu\omega + \nu - \tfrac{3}{2}(\omega_x + \omega\omega_y) = 0. \tag{13}$$

Again the equations allow of geometrical interpretation. (12) stands for the two equivalent dynamical properties, III and IV:

Property III. The locus of the centers of the ∞^1 circles corresponding to the lineal elements at a given point is a conic with that point as a focus.

Property IV. In each direction through a given point O there passes one trajectory which has contact of third order with its circle of curvature. The locus of the centers of the ∞^1 hyperosculating circles, obtained by varying the initial direction, is a conic passing through the given point in the direction of the force.

Property V. Of the curves which pass through a given point in the

[14] Kasner, *Bulletin of the American Mathematical Society*, vol. 14 (1908), p. 356; vol. 36 (1930), p. 51.

direction of the "force" at that point, there is one which has contact of the third order with its circle of curvature; the radius of curvature of this curve is three times the radius of curvature of the "line of force" (i. e., integral curve of the direction assigned to each point by II) passing through the given point.

The analytic statement of this last property is

$$(14) \qquad \lambda\omega^2 + \mu\omega + \nu + \omega_x + \omega\omega_y = 0.$$

This differs only slightly from (13). We can see by following the argument in the reference that (13) changes Property V to this extent: the "three times" in V is replaced by "twice". It is clear that (14) holds in our problem only when $\omega_x + \omega\omega_y = 0$; that is, only when the "lines of force" defined by $y' = \omega(x, y)$ are straight lines.

At this point we recall that the families of conics we are finding are assumed to be curvature trajectories. A function $F(x, y, y')$ therefore exists such that

$$(15_{1,2}) \qquad G = (F_x + F_y y')/F, \qquad H = F_{y'}/F,$$

as we see on comparing (3) and (4). From (15_2) and (11_2) we find that

$$(16) \qquad F = (y' - \omega)^3\, \theta(x, y),$$

where $\theta(x, y)$ is arbitrary. Substituting this expression for F into (15_1), we have

$$(17) \qquad G = [\bar{\theta}_y y'^2 + (\bar{\theta}_x - \omega\bar{\theta}_y - 3\omega_y)y' - (\omega\bar{\theta}_x + 3\omega_x)]/(y' - \omega),$$

where $\bar{\theta} = \log\theta$. We now equate coefficients in (17) and (12), and use (13). We find without difficulty that the existence of θ has the following two consequences:

$$(18) \qquad \omega_x + \omega\omega_y = 0,$$

$$(19) \qquad \lambda_x + [(\nu + 3\omega_x)/\omega]_y = 0.$$

The first of these, as we have said, means that Property V holds. It also means that the denominator of G in (12) divides the numerator; (13) shows this directly. Thus (12) reduces to

$$(20) \qquad G = \lambda y' - \nu/\omega.$$

With (19) we may compare

$$(21) \qquad \lambda_x + [(\nu + \omega_x)/\omega]_y = 0,$$

which is the analytic form of the final one of Kasner's set of properties of dynamical trajectories:

Property VI. When the point O is moved, the associated conic described

in Property IV changes in the following manner. Take any two fixed perpendicular directions for the x direction and the y direction; through O draw lines in these directions meeting the conic again at A and B, respectively. Also construct the normal at O meeting the conic again at N. At A draw a line in the y direction meeting this normal in some point A', and at B draw a line in the x direction meeting the normal in some point B'. The variation property referred to, takes the form

$$(22) \qquad \frac{\partial}{\partial x}\frac{1}{AA'} + \frac{\partial}{\partial y}\frac{1}{BB'} + \frac{\omega\omega_{xy} - \omega_x\omega_y}{3\omega^2} = 0,$$

where AA' and BB' denote distances between points, and where ω denotes the slope of the lines of force relative to the chosen x direction. This is true for any pair of orthogonal directions, and therefore really expresses an intrinsic property of the system of curves. (See diagram in references.)

As before, we find that (19) is interpreted by simply deleting the 3 in the denominator of the last term of (22).[15] Thus (21) and (19) will both hold at once only when $\omega\omega_{xy} - \omega_x\omega_y = 0$; that is, using (18), only when $\omega_{yy} = 0$. But it is easy to show that $\omega_{yy} = 0$ and (18) imply that either $\omega(x, y) = y/x$ (essentially), or $\omega(x, y) = const.$ Hence to establish Property VI we must prove that the field of force is either central or parallel. To do this, however, we must use the third of equations (10).

We therefore set (20) and (11_2) into (10_3). We need only the two highest coefficients of the resulting identity in y'. From these, with the aid of (18) and (19), we deduce the following equations:

$$(23) \qquad \lambda_x + \omega\lambda_y + \omega_y\lambda - \gamma_y - 3\omega_{yy} = 0,$$

$$(24) \qquad \lambda(\gamma + 3\omega_y) - 3\gamma_y = 0,$$

$$(25) \qquad 3\gamma_x + 3\omega\gamma_y - \gamma^2 = 0,$$

where $\gamma = \lambda\omega - \nu/\omega$. ($\gamma$ represents G for the force direction.) Now (24) serves to eliminate λ from (23), assuming $\gamma + 3\omega_y \neq 0$. By means of (25) and (18) we then eliminate derivatives of γ and ω with respect to x. The result factors into

$$(26) \qquad (\gamma + 3\omega_y)^2\,\omega_{yy} = 0.$$

If $\gamma + 3\omega_y = 0$, then differentiating, $\gamma_y + 3\omega_{yy} = 0$; but from (24), $\gamma_y = 0$. Hence in any case $\omega_{yy} = 0$. Property VI therefore does hold.

Since Properties I-VI form a characteristic set for dynamical trajectories, we have proved that in the dynamical case, curvature families of conics must

[15] It is curious that this change produces Property VI of *velocity curves.* Similarly, the change in V gave Property V of *catenaries.* (*Princeton Colloquium,* p. 94.)

be dynamical trajectories; that is, they must be one of the three kinds a), b), and c) in B1. It follows from A2 (since a), b), c) come from central fields) that all three kinds are in fact curvature trajectories. Hence a), b), c) are the answers in the dynamical case.

We observe that the last point could be established independently of A2; for the process of applying the dynamical properties expressed in (4), (11_2), (12_2), (14), and (21) would closely parallel the work above, and would give exactly the same result. In other words, the above proof can be turned into a solution of the first part of Bertrand's problem, to show that fields of force whose trajectories are conics must be central or parallel.

Sectional case. As we do not have a set of properties for sectional families like those for dynamical families, we cannot give a geometrical interpretation for each step in this case. We outline the calculations briefly. A point of interest is that only (10_2) is used, whereas the dynamical case required (10_3) as well.

H is now supposed to be given by (11_1) with $\rho\tau - \sigma^2 \neq 0$. Applying ($15_2$), we determine the form of F in y'; this is found to be the same as in (5) Then from (15_1) we find that G has the same form in y' as the G of the sectional type (2). The expressions for H and G are now substituted into (10_2), and the resulting identity in y' yields a system of four equations from which all the unknowns can be calculated. In this way we obtain for F the following expression:

$$(27)\quad F = \left[\frac{(a_{33}y^2 - 2a_{23}y + a_{22}) - 2(a_{33}xy - a_{23}x - a_{13}y + a_{12})y' + (a_{33}x^2 - 2a_{13}x + a_{11})y'^2}{A_{11}x^2 + 2A_{12}xy + A_{22}y^2 + 2A_{13}x + 2A_{23}y + A_{33}}\right]^{3/2},$$

where $a_{11}, \cdots, a_{33}$ are arbitrary constants except that $|a_{ij}|$, (where $a_{ji} = a_{ij}$), is of rank at least 2, and where $A_{11}, \cdots, A_{33}$ are cofactors in $|a_{ij}|$.

The final step consists in forming the F for an arbitrary proper quadric, (say by means of the formula implied in the proof of A3), and identifying it with (27). Two cases are necessary, depending on whether or not the quadric contains the point at infinity on the z-axis (the center of projection). The identifying requires a theorem on minors in an adjoint determinant; (Bôcher, *Algebra*, p. 31).

The families of conics obtained in this case are therefore d) and e) of B2. The proof of B3 is thus complete.

7. Part of the last derivation applies to a more general problem: to determine all families of conics having Property I. This includes B1, B2,

and B3. The answers here can be converted by duality, since Property I is invariant under correlations.[16]

We give examples to show that the solutions to the proposed problem will add to the set a), b), c), d), e). The following triply infinite families of conics have Property I, but are neither dynamical nor sectional:

f) Let points A, B, C be the vertices of a triangle. A conic through A will cut sides AB and AC again in points M and N. The family consists of the ∞^3 conics through A for which the tangents at M and N meet on BC.

g) Let the pencil of lines through a fixed point A be in projective correspondence with the points of a fixed line l through A, in such a way that l corresponds to A. A conic through A will be tangent at A to a line u of the pencil, and will cut l again in a point X. The family consists of the ∞^3 conics through A for which u and X correspond.

An alternative form of Property I states that the ∞^1 osculating parabolas at a given lineal element have concurrent directrices.[17] Family f) may be considered an obvious construction from this form of Property I; for if we dualize f) and specialize the triangle properly, we obtain the family of ∞^3 parabolas whose directrices pass through a given point. Family g) is a limiting case of f). Both f) and g) arise in the sectional case of the problem. They are the only new solutions for which one set of "asymptotic curves" is a pencil of straight lines.

In the dynamical case, it can be shown that there are no new solutions, under the assumption that the "lines of force" are a family of straight lines. An equivalent statement is this: Properties I-V suffice to define a), b) and c) as the only families of conics which are dynamical trajectories. That is, in solving Bertrand's problem it is not necessary to impose all the six geometrical properties of dynamical trajectories, for the sixth property becomes a consequence of the other five.

An example of a new family in the dynamical case is given by the ∞^3 parabolas with directrices through a fixed point. The "lines of force" are the ∞^1 circles concentric about the point. This family has Properties I, II, III, IV, and also VI, but not V. It is therefore as nearly dynamical as any non-dynamical family of conics can be, in the sense that it has a maximum number of the six properties without having all of them. It might be termed the next thing to a Bertrand family.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

[16] If we apply Legendre's transformation (polarity with respect to the parabola $x^2 - 2y = 0$) to equations (10), and interchange G with $-H$, we obtain the same equations in reverse order. This expresses the fact that the dual of a Property I family of conics is such a family. (See *Princeton Colloquium*, p. 78.)

[17] *Princeton Colloquium*, p. 12.

SOME INTERPRETATIONS OF ABSTRACT LINEAR DEPENDENCE IN TERMS OF PROJECTIVE GEOMETRY.[1]

By Saunders MacLane.

1. *Introduction.* The abstract theory of linear dependence, in the form recently developed by Whitney,[2] is closely related to the study of projective configurations. For any matroid (that is, any finite system of elements for which a suitably restricted notion of "linear dependence" is given) can be interpreted as a schematic geometric figure. Such a schematic figure, like a schematic configuration, is composed of a number of points, lines, planes, etc., with certain combinatorially defined incidences. The problem of representing a matroid by a matrix then becomes simply the problem of realizing a schematic figure by some geometric figure—and the impossibility of always finding such a representation turns out to be a simple consequence of Pascal's theorem! Even when such representation is possible, it depends essentially upon the field from which the elements of the representing matrix are taken. However, only algebraic fields need be used, and hence arises a connection between certain matroids and the algebraic fields in which they can be best represented.

Matroids will be defined by axioms on "rank," as in Whitney's paper. Without loss of generality we can also assume the following two axioms:

R_4: The rank of a single element is always 1.

R_5: The rank of a pair of elements is always 2.

For example, an element e which does not satisfy R_4 may be dropped from or added to a matroid M without otherwise altering the structure of M. These two axioms are equivalent to the following conditions on "bases":

B_3: Every element belongs to at least one base.

B_4: There is no pair of elements (e_1, e_2) such that every base containing e_1 remains a base when e_1 is replaced by e_2.

These conditions on M are in turn equivalent to the following conditions on the dual matroid M^*:

C^*_3: Every element is omitted from at least one circuit complement.

[1] Presented to the American Mathematical Society, December 28, 1934.

[2] H. Whitney, "On the abstract properties of linear dependence," *American Journal of Mathematics*, vol. 57 (1935), pp. 509-533.

C^*_4: For every pair of elements (e_1, e_2) there is a circuit complement containing e_1 but not e_2.

2. *Schematic geometric figures.* A rectilinear plane figure consists of a number of points and of all the lines joining these points in pairs. The combinatorial structure of such a figure can be specified by giving for each line L the set of all those points of the figure which lie on L. These sets satisfy the following axioms:

F_1: Any pair of points belongs to one and only one line.
F_2: Every line contains at least two points.
F_3: No line contains all the points.
F_4: There are at least two points.

A system consisting of a finite number of "points" and certain sets of these points, called "lines," and satisfying these axioms will be called a *schematic plane figure*; it may or may not correspond to some actual figure. In the same way a *schematic space figure* would involve "points," "lines," and "planes," satisfying F_1 to F_4 and the following axioms:

S_1: Every triple of points belonging to no line belongs to one and only one plane.
S_2: Every plane contains three points not on a line.
S_3: No plane contains all the points.
S_4: If a plane contains two points of a line, it contains all the points of that line.

The definition of a schematic n-dimensional figure is similar; it involves "points" and "k-dimensional planes" for $k = 1, \cdots, n-1$.

The equivalence of schematic figures and matroids may be formulated as follows:

THEOREM 1. *Every schematic n-dimensional figure is a matroid of rank $n+1$ if the rank of a set of points A is defined as the smallest r such that all the points of A are contained in some $(r-1)$-plane. Conversely, every matroid of rank $n+1$ becomes a schematic n-dimensional figure if the k-planes are taken as maximal sets of elements of rank $k+1$. This translation sets up a one-one correspondence between matroids and schematic figures.*[3]

From this theorem it follows that a schematic n-dimensional figure is

[3] The conditions R_4 and R_5 of § 1 are necessary here to exclude the geometrically meaningless cases of a point of dimension -1 or of two coincident points.

determined once the corresponding matroid M or its dual M^* is given. By Whitney's results, this dual M^* is completely determined by its circuit complements. The circuit complements in M^* correspond to maximal sets of rank $r(M)-1$ in M. Hence

THEOREM 2. *A schematic n-dimensional figure is completely determined if its $(n-1)$-planes are given. If a set of "points" and certain subsets of this set are given, these subsets will be the $(n-1)$-planes of some figure if and only if they are the circuit complements of a matroid M^*; that is, if and only if these subsets satisfy the above axioms C^*_3 and C^*_4, while their complements satisfy Whitney's axioms C_1 and C_2 for circuits.*

These results also show that matroids form a direct generalization of schematic configurations. A *schematic plane configuration*[4] $p_\gamma g_\pi$ consists of p "points" and g "lines," with each point on γ lines and each line on π points. Such a configuration becomes a schematic figure in the above sense if those pairs of points not already joined by lines are joined by new "diagonal" lines. Similar transformations are possible for space configurations.

3. *Matrix representations of matroids.* The columns of a matrix stand in relations of rank and thus form a matroid. The question whether every matroid can be represented in this way by a matrix is clarified by the equivalence of matroids and schematic figures. Thus Whitney has constructed a matroid of rank 3 which cannot be represented as a matrix. This matroid has 9 elements $1, 2, \cdots, 9$ and the following 20 circuit complements:

$$\begin{gathered}712,\ 814,\ 923,\ 734,\ 836,\ 945,\ 756,\ 825;\\ 16,\ 19,\ 69,\ 13,\ 15,\ 24,\ 26,\ 35,\ 46,\ 78,\ 79,\ 89.\end{gathered}$$

Any attempt to represent this matroid yields a figure in which the lines 16, 19, and 69 coalesce into one line 169. A geometric representation reveals at once that this is simply Pascal's theorem for the hexagon 723845 inscribed in the degenerate conic composed of the two lines 743 and 825. The points 1, 6, and 9 are the intersections of opposite sides of the hexagon. In exactly the same way the theorem of Desargues may be used to construct a matroid with ten elements which has no matrix representation. Furthermore, the matroid arising from Pascal's theorem can be generalized to the case of $2m+3$ elements, which we denote by $1, 2, \cdots, 2m, \alpha, \beta, \gamma$. The circuit complements are:

[4] F. Levi, *Geometrische Konfigurationen*, p. 4.

$$12\alpha,\ 34\alpha,\ \cdots,\ (2m-3,\ 2m-2,\ \alpha),\ (2m-1,\ 2m,\ \alpha),$$
$$14\beta,\ 36\beta,\ \cdots,\ (2m-3,\ 2m,\ \beta),\ (2m-1,\ 2,\ \beta),$$
$$23\gamma,\ 45\gamma,\ \cdots,\ (2m-2,\ 2m-1,\ \gamma).$$

together with all the pairs of elements not included in one of these triples. No matrix representation is possible, for any attempt to construct one yields a matroid with the additional circuit complement $(2m, 1, \gamma)$.

These matroids fail to be matrices because of the presence of too few circuit complements. Failure is also possible for the opposite reason. Thus the plane figures (matroids) formed by finite projective geometries [5] can be represented only by matrices of elements from a finite field. Another important special case of the matrix representation of matroids is the problem of constructing a geometric realization for schematic plane configurations. Here it is well-known that a configuration (p_γ, g_π) cannot in general be realized if [6]

$$2(p+g)-p\gamma-8<0.$$

The use of geometric figures also simplifies the investigation of the conditions for the representability of individual matroids. Thus for a matroid M of rank 3 we need only find three homogeneous coördinates for each element (point) of the matroid, such that when three points lie on a line (i. e., are contained in a circuit complement of the dual matroid), then the determinant of the corresponding coördinates is zero, and conversely. This application of the usual theorems of analytic geometry can replace Whitney's Theorem 32.

4. *Representation in finite algebraic fields.* The configuration of eight elements which can be represented in the complex but not in the real plane [7] suggests that the representability of a matroid depends essentially on the field from which the elements of the representing matrix are taken. Another similar example can be constructed for the field $R(2^{1/2})$, where R is the field of rational numbers. We need only take a point with coördinates $(1, 2^{1/2}, 0)$ and carry out the constructions in the von Staudt algebra of throws corresponding to

$$(2^{1/2})(2^{1/2})=1+1.$$

The resulting figure (matroid) consists of 11 points, $1, 2, \cdots, 9, 0, \alpha$, the following sets of points being collinear:

$$1279\alpha,\ 2356,\ 1380,\ 248,\ 347,\ 578,\ 549,\ 690,\ 50\alpha,\ 68\alpha.$$

---

[5] Veblen and Young, *Projective Geometry*, vol. I, p. 3 and p. 201.

[6] E. Steinitz, *Encyklopädie der mathematischen Wiss.*, III A B 5a, p. 485.

[7] F. Levi, *loc. cit.*, pp. 98-102.

Any attempt to represent this matroid by a matrix leads to a matrix whose elements generate a field containing $R(2^{\frac{1}{2}})$. This matroid is thus a sort of geometric analog of the irreducible equation for $2^{\frac{1}{2}}$. Generalization yields:

THEOREM 3. *Let K be a finite algebraic field over the field of rational numbers. Then there exists a matroid M of rank 3 which can be represented by a matrix with elements in K, while any other representation of M by a matrix with elements in a number-field K_1 requires $K_1 \supset K$.*

Such finite fields are sufficient for the representation of all representable matroids, in the following sense.

THEOREM 4. *Let the matroid M be representable by a matrix of complex numbers. Then M can also be represented by a matrix with elements from an algebraic field of finite degree.*

For let the matroid M have rank n and consist of p points, and let these points be assigned the indeterminate homogeneous coördinates a_{ij}, for $i = 1, \cdots, n$; $j = 1, \cdots, p$. Each circuit complement of the dual of M requires the vanishing of a number of determinants of the a_{ij}, and thus corresponds to a number of algebraic equations for these quantities. The set of all values of the a_{ij} giving at least the required circuit complements thus constitutes an algebraic manifold N_1 in the np-dimensional space of all coördinates a_{ij}. The set of those coördinates yielding additional undesired circuit complements forms another manifold N_2. Since the original matrix was representable, there exists a point of $N_1 - N_2$. The parametric representation of irreducible algebraic manifolds [8] thus makes possible the construction of a point with algebraic coördinates in $N_1 - N_2$.

HARVARD UNIVERSITY.

---

[8] B. L. van der Waerden, *Moderne Algebra*, vol. 2, p. 51 ff.

WARING THEOREMS OF NEW TYPE.

By L. E. Dickson.

1. Denote $ax^3 + by^3$ by (a, b) and similarly for several variables, which we restrict to integral values $\geqq 0$. We shall prove

Theorem 1. *Every integer $\geqq 0$ is a sum of three values of any of the forms* $(1, 1, j)$, $j = 1, 2, 5$; $(1, 2, k)$, $k = 2, \cdots, 6$.

For $(1, 1, 1)$, this is Waring's theorem on nine cubes.

Theorem 2. *Every integer $\geqq 0$ is a sum of five values of* $(1, 2)$. *All integers* $> 11^{12}$ (19.006868) *are sums of five values of* $(1, 5)$.

In Theorem 2 we may take zero as one of the ten cubes.

Call a form universal if it represents all integers $\geqq 0$. Theorem 5 gives 6344 universal forms each a sum of nine products of a cube by a positive integer. Certain of them yield Theorem 1.

Our method is a direct generalization of the classic one for nine cubes.[1] For some of our forms the auxiliary work in Lemma 3 is much simpler than that for nine cubes. The problem is under investigation by different methods.

2. *The four lemmas.*

Lemma 1. *If p is a prime $\equiv 2 \pmod 3$, every integer not divisible by p is congruent to a cube modulo p^n.*

This well known Lemma (*Transactions, loc. cit.*) implies

Lemma 2. *If $p \equiv 2 \pmod 3$, and if l is not divisible by the prime p, every integer not divisible by p is congruent modulo p^n to the product of a cube by l.*

Consider the forms

$$\begin{array}{lll} A = x^2 + y^2 + z^2, & B = x^2 + y^2 + 2z^2, & C = x^2 + y^2 + 5z^2, \\ D = x^2 + 2y^2 + 2z^2, & E = x^2 + 2y^2 + 3z^2, & F = x^2 + 2y^2 + 4z^2, \\ G = x^2 + 2y^2 + 5z^2, & H = x^2 + 2y^2 + 6z^2. & \end{array}$$

[1] Dickson, *Transactions of the American Mathematical Society*, vol. 30 (1928), pp. 1-18. The discussion on pp. 7-13 becomes simpler by our new Lemma 4 since we now obtain the final C at once without reducing a preliminary C. As on p. 14, we readily extend our Lemma 3 to further values of P.

LEMMA 3. *Let P and e be given integers $\geqq 0$. If $P \equiv 5 \pmod{16}$, every integer $\geqq P^e \cdot 11^3$ is represented by $P^e\gamma^3 + 6f$, where $\gamma \geqq 0$ and f is B, E, or F. If $P \equiv 5 \pmod{48}$, every integer $\geqq P^e \cdot 22^3$ is represented by $P^e\gamma^3 + 6f$, where $\gamma \geqq 0$ and f is A, D, or H. If $P \equiv 11 \pmod{48}$, every integer $\geqq P^e \cdot 23^3$ is represented by $P^e\gamma^3 + 6f$, $\gamma \geqq 0$, $f = A$ or C. If $P \equiv 11 \pmod{50}$, every integer $\geqq P^e \cdot 11^3$ is represented by $P^e\gamma^3 + 6G$ with $\gamma \geqq 0$.*

The proof is simplest for B, E, F, or G. Elsewhere [2] I proved that E represents exclusively all positive integers not of the form $4^k(16r + 10)$. Hence E represents every positive integer congruent modulo 16 to

$$(1) \qquad 1,\ 3,\ 4,\ 5,\ 6,\ 7,\ 9,\ 11,\ 12,\ 13,\ 14,\ 15,$$

or 2, which we omit. Let (γ, μ) denote the least positive residue of $\gamma^3 + 6\mu$ modulo 96. The 72 distinct numbers (γ, μ), with $0 \leqq \gamma \leqq 5$ and μ ranging over (1), together with

$$\begin{gathered}(6,4) = 48,\ (6,6) = 60,\ (6,12) = 0,\ (6,14) = 12,\ (7,1) = 61,\\ (7,7) = 1,\ (7,9) = 13,\ (7,15) = 49,\ (8,4) = 56,\ (8,6) = 68,\\ (8,12) = 8,\ (8,14) = 20,\ (9,3) = 75,\ (9,5) = 87,\ (9,11) = 27,\\ (9,13) = 39,\ (10,4) = 64,\ (10,6) = 76,\ (10,12) = 16,\ (10,14) = 28,\\ (11,1) = 89,\ (11,7) = 29,\ (11,9) = 41,\ (11,15) = 77,\end{gathered}$$

are found to be a rearrangement of $0, 1, \cdots, 95$. Since the product of the numbers (1) by 5 (and hence by P) are congruent modulo 16 to the same 12 numbers permuted, we conclude that, if n is any integer,

$$(2) \qquad n \equiv P^e\gamma^3 + 6\mu \pmod{96}$$

has integral solutions with $0 \leqq \gamma \leqq 11$, μ in (1). Thus there is an integer q for which

$$(3) \qquad n = P^e\gamma^3 + 6\mu + 96q = P^e\gamma^3 + 6m, \qquad m = \mu + 16q.$$

Let $n \geqq P^e \cdot 11^3$. Then $n \geqq P^e\gamma^3$, $m \geqq 0$, whence m is represented by E.

Proof for B. The form B represents exclusively all positive integers not of the form $4^k(16r + 14)$. Hence B represents all congruent modulo 16 to

$$(4) \qquad 1,\ 2,\ 3,\ 4,\ 5,\ 7,\ 9,\ 10,\ 11,\ 12,\ 13,\ 15,$$

or 6. The 72 residues (γ, μ), with $0 \leqq \gamma \leqq 5$ and μ in (4), together with

[2] *Bulletin of the American Mathematical Society*, vol. 33 (1927), pp. 63-70.

$(6,2)=36$, $(6,4)=48$, $(6,9)=78$, $(6,10)=84$, $(6,12)=0$,
$(7,5)=85$, $(7,7)=1$, $(7,13)=37$, $(7,15)=49$, $(8,2)=44$,
$(8,4)=56$, $(8,10)=92$, $(8,12)=8$, $(9,1)=63$, $(9,3)=75$,
$(9,9)=15$, $(9,11)=27$, $(10,2)=52$, $(10,4)=64$, $(10,10)=4$,
$(10,12)=16$, $(11,5)=17$, $(11,7)=29$, $(11,13)=65$, $(11,15)=77$,

form a rearrangement of $0,1,\cdots,95$. The products of (4) by 5 are congruent modulo 16 to the same 12 numbers rearranged. Hence (2) is solvable with $0 \leqq \gamma \leqq 11$, μ in (4). We have (3) with m represented by B. This proof applies unchanged also to F.

Proof for G. By *Bulletin, loc. cit.*, G represents exclusively all positive integers $\neq 25^k(25r \pm 10)$ and hence represents all $\equiv \mu \pmod{50}$, $\mu = 1,\cdots,9,\quad 11,\cdots,14,\quad 16,\cdots,24,\quad 26,\cdots,34,\quad 36,\cdots,39,\quad 41,\cdots,49$. Since $11 \times 5 \equiv 5$, multiplication yields $11j \equiv j \pmod{50}$ for $j = 5, 20, 30, 45$, i. e., for all the values of μ which are multiples of 5. When x ranges over the integers prime to 5 and < 50, the same is true of the residues of $11x$ modulo 50. Hence μ and 11μ take the same residues modulo 50. For $0 \leqq \gamma \leqq 11$, the residues modulo 300 of $\gamma^3 + 6\mu$ are found to be $0, 1, \cdots, 299$. Multiplication by $P \equiv 11 \pmod{50}$ shows that $n \equiv P^e\gamma^3 + 6\mu \pmod{300}$ is solvable for every n with $0 \leqq \gamma \leqq 11$.

Proof for H. The form H represents[3] exclusively all positive integers not of the form $4^k(8s+5)$. Hence H represents all $\equiv$

$$2,\ 3,\ 6,\ 7,\ 8,\ 10,\ 11,\ 12,\ 14,\ 15, \tag{5}$$

or 1 or 9 (mod 16). Employ $0 \leqq \gamma \leqq 5$, all μ in (5); $\mu = 7, 11$, $\gamma = 7, 11, 15$; $\mu = 11, 15$, $\gamma = 9, 13, 17$; $\mu = 8$, $\gamma = 7, 9, 10, 11, 18$; $\mu = 12$, $\gamma = 6, \cdots, 9, 11, 14, 22$. The resulting residues (γ, μ) form a permutation of the numbers $0, \cdots, 95$ except $6 + 8r$ $(r = 0, \cdots, 11)$. The latter are the values of (γ, μ) for $\mu = 1, 9$, $\gamma = 0, 2, 4, 6, 8, 10$. Hence, for $e = 0$, (2) is solvable with $0 \leqq \gamma \leqq 22$ and $\mu = 1, 9$ or one of (5). In $5(6+8r) = 6 + 8\rho$, $\rho = 3 + 5r$ ranges with r over a complete set of residues modulo 12. Hence multiplication by 5 of the 12 numbers $6 + 8r$ or of the 10 numbers (5) merely permutes them modulo 96. Evidently the same holds when 5 is replaced by P. The final step of the proof for H is the same as for A (*Transactions, loc. cit.*). That proof for A applies unchanged to D since A and D each represent exclusively all positive integers $\neq 4^k(8s+7)$.

[3] B. W. Jones, *Chicago Dissertation* (unpublished), 1928, p. 72.

Proof for C. The form C represents[4] exclusively all positive integers $\neq 4^k(8s+3)$. Hence C represents all $\equiv 1, 4, 9$ or

(6) $$2,\ 5,\ 6,\ 7,\ 8,\ 10,\ 13,\ 14,\ 15,$$

modulo 16. Employ $0 \leqq \gamma \leqq 5$, all μ in (6); $\mu = 5, 7, 8, 13, 15$, $\gamma = 6, 8, 10$; $\mu = 6, 14$, $\gamma = 9, 13, 17, 21$; $\mu = 2, 7, 8, 10$, $\gamma = 7, 11$; $\mu = 7, 15$, $\gamma = 9$; $\mu = 2$, $\gamma = 15, 19, 23$; $\mu = 8$, $\gamma = 14, 18, 22$. The resulting residues (γ, μ) give $0, \cdots, 95$ except $0, 32, 64$. The proof for C is now completed as for A in (*Transactions, loc. cit.*, p. 9), since the products of (6) by 11 are congruent modulo 16 to the same nine numbers permuted.

LEMMA 4. *Given R and the positive numbers s and t such that*

(7) $$s \geqq t + t/9^3, \qquad t \leqq R \leqq s,$$

we can find an integer $i \geqq 0$ satisfying

(8) $$R \leqq s - ti^3 < R + K, \qquad K = 3[t(s-t)^2]^{1/3}.$$

If $s < R + K$, (8) holds for $i = 0$. Henceforth, let $s \geqq R + K$. Determine a real number r so that $s - tr^3 = R$. By (7_1),

$$tr^3 \geqq K \geqq t/27, \quad r \geqq 1/3.$$

We may write $r = i + f$, where $0 \leqq f < 1$ and i is an integer $\geqq 0$. Since $i \leqq r$, $-r^3 \leqq -i^3$ and $R \leqq s - ti^3$, as desired in (8). Next,

$$s - ti^3 - R = s - t(r-f)^3 - s + tr^3 = tw,$$
$$w = r^3 - (r-f)^3 = 3r^2f - f^2(3r - f) < 3r^2f < 3r^2,$$

since $3r - f > 0$. By $t \leqq R$, $tr^3 \leqq s - t$, whence

$$3tr^2 \leqq K, \ s - ti^3 - R = tw < 3tr^2 \leqq K.$$

3. *General theory.* When $C > 0$ and $p > 1$, Cp^{3x} increases indefinitely with x. To any $s - t > 0$ therefore corresponds an integer $n \geqq 0$ for which

(9) $$Cp^{3n} \leqq s - t < Cp^{3(n+1)}.$$

THEOREM 3. *Let s, t, l, h_1, h_2, h_3 be given positive integers for which $s \geqq t + t/9^3$. Choose a prime p, $p \equiv 2 \pmod 3$, which divides neither t nor l, such that, as a composite of the cases in Lemma 3 with $P = p$, every integer $\geqq p^e g^3$ is represented by $p^e\gamma^3 + 6f$ for every integer $e \geqq 0$, where $\gamma \geqq 0$ and*

[4] Dickson, *Annals of Mathematics* (2), vol. 28 (1927), pp. 340-341.

$$(10) \qquad f = h_1x_1^2 + h_2x_2^2 + h_3x_3^2.$$

Restrict e and n to integral values for which

$$(11) \qquad e = 0, 1, \text{ or } 2, \qquad e + n \equiv 0 \ (mod\ 3).$$

Choose ν and C so that, when $n \geqq \nu$ (see § 4),

$$(12) \qquad (C - l - w)p^{2n-e} \geqq g^3,\ Cp^{3n} + 2k \leqq (6 + w)p^{3n},\ Cp^{3n} \geqq t,$$

where

$$(13) \qquad w = 2(h_1 + h_2 + h_3),\ k = 3(tC^2)^{1/3}p^{2n+2}.$$

Determine n by (9). *Then if $s \geqq t + Cp^{3\nu}$, there exist nine integers a, b, c, y_i, z_i $(i = 1, 2, 3)$, each $\geqq 0$, such that*

$$(14) \qquad s = ta^3 + lb^3 + c^3 + \sum_{i=1}^{3} h_i(y_i^3 + z_i^3).$$

The final inequality (9) gives $K < k$ for K in (8).

First, let $Cp^{3n} + 2k \leqq s$. Then Cp^{3n} and $Cp^{3n} + k$ are values of R satisfying (7_2). Hence by (8) there exist integers I and J, each $\geqq 0$, such that

$$Cp^{3n} \leqq s - tI^3 < Cp^{3n} + k,$$
$$Cp^{3n} + k \leqq s - tJ^3 < Cp^{3n} + 2k.$$

Since these intervals contain no common number, $I \neq J$. Hence there exist two distinct integral values of i which satisfy

$$(15) \qquad Cp^{3n} \leqq s - ti^3 < Cp^{3n} + 2k, \qquad i \geqq 0.$$

Second, let $Cp^{3n} + 2k > s$. By (9_1), (15) holds when $i = 0$ or 1. Hence in both cases there exist two distinct integers and hence two consecutive integers $j - 1$ and j which are both values of i satisfying (15). At least one of the integers $s - t(j-1)^3$ and $s - tj^3$ is not divisible by p. In fact, their difference is the product of t by $3j^2 - 3j + 1$. If $p = 2$, this is not divisible by p. If $p > 2$,

$$12(3j^2 - 3j + 1) = (6j - 3)^2 + 3$$

is not divisible by p since the reciprocity law gives

$$(-3 \mid p)(p \mid 3) = 1, \quad (p \mid 3) = (2 \mid 3) = -1, \quad (-3 \mid p) = -1.$$

Hence there exists an integer $a \geqq 0$ such that (15) holds when $i = a$, and such that $s - ta^3$ is not divisible by p. By Lemma 2, $s - ta^3 \equiv l\beta^3 \pmod{p^n}$, where β is not divisible by p. Evidently $\beta \equiv b$, where

$$(16) \qquad s - ta^3 = lb^3 + p^n M, \quad 0 < b < p^n.$$

By (12_2) (16) and (15) with $i = a$,

$$Cp^{3n} \leqq lb^3 + p^n M < (6 + w)p^{3n}, \quad (C - l)p^{3n} < Cp^{3n} - lb^3,$$
$$(C - l)p^{3n} < p^n M < (6 + w)p^{3n}.$$

Cancel p^n and write $M = N + wp^{2n}$. Thus

$$(17) \qquad (C - l - w)p^{2n} < N < 6p^{2n},$$

$$(18) \qquad s = ta^3 + lb^3 + p^n(N + wp^{2n}).$$

By an assumption in Theorem 3, N is represented by $p^e\gamma^3 + 6f$ if $N \geqq p^e g^3$ and hence by (17_1) if (12_1) holds. By (11), $p^{n+e}\gamma^3$ is the cube of an integer $c \geqq 0$. Thus

$$(19) \qquad p^n N = c^3 + 6fp^n,$$

$$(20) \qquad s = ta^3 + lb^3 + c^3 + p^n(wp^{2n} + 6f).$$

Since $w = 2\Sigma h_i$,

$$(21) \qquad \sum_{i=1}^{3} h_i\{(p^n + x_i)^3 + (p^n - x_i)^3\} = wp^{3n} + 6fp^n.$$

Evidently (20) now gives (14). If $x_i > p^n$ for a certain i, then $f > p^{2n}$ by (10), and $N > 6p^{2n}$ by (19), contrary to (17). Hence each of our nine cubes is $\geqq 0$.

4. *Conditions* (12). For[5] $n \geqq 3$, the minimum of $2n - e$ is 6, and (12_1) holds if

$$(22) \qquad C - l - w \geqq (g/p^2)^3.$$

Henceforth let $n \geqq 4$. Then $1/p^{n-2} \leqq 1/p^2$ and (12_2) holds if

$$(23) \qquad C + (tC^2)^{1/3}(6/p^2) \leqq 6 + w.$$

For example, consider D. Here $p = 5$, $g = 22$, $w = 10$. First, let $l \geqq 4$. Then $C > 14\frac{4}{5}$ by (22), $C^2 > 215$, $C^{2/3} > 5.99$ and (23) fails. When $l = 3$, $C \geqq 13.68147$ and (23) gives $t \leqq 4$.

In this way we obtain the complete solution of (12) when $n \geqq 4$:

[5] Let $n = 3$. We have (23) with p^2 replaced by p. This case is excluded for forms B, D, E, F, G. But for C, it arises only when $l = 2$, $t = 1$; $l = 1$, $t = 1, 2, 3$. For H, there are many sets l, t. Any set with $\nu = 3$ is evidently a set with $\nu = 4$. We obtain universal theorems for all the latter sets. Hence we ignore those special sets which are also sets with $\nu = 3$, in spite of the lowering of the constant $Cp^{3\nu}$.

B, $p=5$.	$l=4$, $t\leqq 3$;	$l=3$, $t\leqq 14$;	$l=2$, $t\leqq 42$;
		$l=1$, $t\leqq 104$.	
C, $p=11$.	$l=5$, $t\leqq 22$;	$l=4$, $t\leqq 200$;	$l=3$, $t\leqq 760$;
		$l=2$, $t\leqq 2038$;	$l=1$, $t\leqq 4533$.
D, $p=5$.	$l=3$, $t\leqq 4$;	$l=2$, $t\leqq 16$;	$l=1$, $t\leqq 42$.
E, $p=5$.	$l=4$, $t=1$;	$l=3$, $t\leqq 7$;	$l=2$, $t\leqq 21$;
		$l=1$, $t\leqq 50$.	
F, $p=5$.	$l=4$, $t=1$;	$l=3$, $t\leqq 6$;	$l=2$, $t\leqq 16$;
		$l=1$, $t\leqq 37$.	
G, $p=11$.	$l=5$, $t\leqq 18$;	$l=4$, $t\leqq 163$;	$l=3$, $t\leqq 612$;
		$l=2$, $t\leqq 1619$;	$l=1$, $t\leqq 3545$.
H, $p=5$.	$l=3$, $t=1$;	$l=2$, $t\leqq 6$;	$l=1$, $t\leqq 15$.

THEOREM 4. *For all these sets l, t such that t is not divisible by p, and for the least C determined by* (22), *conditions* (12) *hold, so that the final statement in Theorem* 3 *is proved.*

5. *Universal theorems.* We employ a lemma whose proof (*Transactions*, p. 4) is very similar to that of our Lemma 4.

LEMMA 5. *Given the positive numbers s and t and a number T for which $0\leqq T\leqq s$, $t\leqq 9^3 s$, we can find an integer $i\geqq 0$ satisfying*

$$T\leqq s-ti^3<T+3(ts^2)^{1/3}. \tag{24}$$

Consider the unfavorable case $G=x^2+2y^2+5z^2$, $p=11$, $l=1$, $t=\tau$, $\tau\leqq 3545$. By (22), the minimum C is 17.00075. The form[6] $J=(11122)$ represents all positive integers $<40{,}000$. For $3s^{2/3}=40{,}000$, Lemma 5 with $T=0$, $t=1$, shows that $s=i^3+J$ if $\log s=6.1874082$. Apply Lemma 5 with $T=0$, $t=5$, and s replaced by σ. Hence if $3(5\sigma^2)^{1/3}=s$, viz., $\log\sigma=8.2159453$, all positive integers $\leqq\sigma$ are represented by $(1_4, 2_2, 5)$. Repeat and take $3(5S^2)^{1/3}=\sigma$, whence $\log S=11.2587510$. Hence all positive integers $\leqq S$ are represented by $1_4, 2_2, 5_2)$. Finally, apply Lemma 5 with $T=0$, $t=\tau$, s replaced by v. Take $3(\tau v^2)^{1/3}=S$, whence $\log v\geqq 14.3976364$. Hence all positive integers $\leqq v$ are represented by $L=(1_4, 2_2, 5_2, \tau)$. By §§ 3, 4, all integers $\geqq K=\tau+Cp^{12}$ are represented by L. Since $\log K<13.7272$, $K<v$. The two results prove that all positive integers are represented by L.

Evidently we obtain a like conclusion if we start with B, D, E, F, or H instead of G, since ascent again begins with J.

For forms $(1_5, 5_2, t, l)$ obtained from C, we take $l=1$ and begin ascent

[6] *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 720.

from (1_6); but we can make only three ascents (instead of four as heretofore). The integers $N < 12{,}000$ which require 7, 8, or 9 cubes are listed by Jacobi.[7] It is readily verified that $N - 5$ is a sum of six cubes unless $N = 47, 111, 300$. But $111 - 5 \cdot 2^3 = 71$ and $300 - 5 \cdot 2^3 = 260$ are sums of five cubes; while $47 - 5 - 5$ is a sum of four. The table by Von Sterneck[8] shows that each integer > 8042 and $< 40{,}000$ is a sum of six cubes. Hence all integers $< 40{,}000$, except only 47, are represented by $(1_6, 5)$ while all are represented by $(1_6, 5_2)$. Taking $t = 5$ and making three ascents, we find that $(1_6, 5_3)$ represents all integers $< s$ if $\log s = 10.143227$. Our least C is 15.006868; for it, $\log (t + 11^{12}C) = 13.6730025 = L$.

But by use of Von Sterneck's table to 40,000, we can show that also all integers from 40,000 to 700,000 are sums of six cubes. The three ascents now lead to s, with $\log s > L$.

THEOREM 5. *The following 6344 forms are universal:* $(1_5, 2_2, t, l)$, $(1_3, 2_4, t, l)$, $(1_3, 2_2, 3_2, t, l)$, $(1_3, 2_2, 4_2, t, l)$, $(1_3, 2_2, 5_2, t, l)$, $(1_3, 2_2, 6_2, t, l)$, *where* t, l *have the values under* B, D, E, F, G, H, *respectively in* § 4. *Also,* $(1_6, 5_3)$.

6. There exist infinitely many primes $p \equiv 2 \pmod 3$ such that p is any of the four forms P in Lemma 3. If $0 < q \leqq 1$, take $p^6 \geqq g^3/q$. Then (22) holds if $C = l + w + q$. Evidently (23) fails if $l \geqq 6$, but holds if $l \leqq 5$ when p is sufficiently large. The same is true when $\nu = 3$ whence p^2 is replaced by p in (23). Hence the general theory in § 3 applies when $\nu = 3$ or 4, $1 \leqq l \leqq 5$, t arbitrary, and p sufficiently large.

Lemma 3 fails for the majority of forms f. For example, $x^2 + y^2 + 3z^2$ represents all positive integers $\neq 9^k(9r + 6)$; but $9 \equiv \gamma^3 + 6\mu \pmod{27}$ requires $\mu \equiv 6 \pmod 9$. Also, $x^2 + y^2 + 6z^2$ represents all $\neq 9^k(9r + 3)$, while $18 \equiv \gamma^3 + 6\mu \pmod{27}$ requires $\mu \equiv 3 \pmod 9$. Next, $x^2 + 3y^2 + 3z^2$ represents all $\neq 9^k(3r + 2)$, while $3 \equiv \gamma^3 + 6\mu \pmod 9$ requires $\mu \equiv 2 \pmod 3$. Finally, $x^2 + y^2 + 4z^2$ represents all except $8r + 3$ and $4^k(8r + 7)$, while $2 \equiv \gamma^3 + 6\mu \pmod 8$ requires $\mu \equiv 3 \pmod 4$.

Report will be made later on the result of using various primes p instead of our single p, on forms involving only eight cubes, and on the generalization of my papers on sums of values of cubic polynomials to sums of their products by arbitrary integers.

THE UNIVERSITY OF CHICAGO.

---

[7] Werke, VI, p. 323; *Journal für Mathematik*, vol. 42 (1851), p. 41.

[8] *Akademie Wissenschaften, Wien, Sitzungsberichte*, vol. 112, IIa (1903), pp. 1627-1666.

FUNDAMENTAL SYSTEMS OF UNITS IN NORMAL FIELDS.

By Marie J. Weiss.

Recently Professor Latimer proved some interesting theorems [1] on the fundamental systems of units in algebraic fields cyclic with respect to the rational field. Here some of these theorems are extended to non-cyclic algebraic fields by a consideration of the integral group ring as a ring of operators for the group of units in the field.[2]

Let Ω be a normal extension of degree g of the rational field and G the Galois group of Ω. From the Galois theory it is known that the maximum real subfield Λ of Ω belongs to a subgroup H of order one or two according as Ω is real or imaginary. In either case denote the degree of Λ by $r+1$. In a normal field a unit η_1 can always be chosen in Λ such that η_1 and its conjugates generate a subgroup M of finite index m in the group of all units U of Ω.[3] Every r of the conjugates form a system of independent units in Ω.

The quotient group U/R, where R is the group of roots of unity in Ω, is invariant under G. Since M is generated by a complete set of conjugates under G, the quotient group M/R', R' being the subgroup of roots of unity found in M, is also invariant under G. When these two quotient groups have been written additively R and R' correspond to zero, and throughout the paper we shall so write them. Note that the classes of U/R containing units which form a fundamental system of units in Ω form an independent basis for U/R, while any r of those classes of M/R' which contain one of the conjugate units $\eta_1, \eta_2, \cdots, \eta_{r+1}$ form an independent basis for M/R'.

If ζ is any unit of Ω and S any element of G, by $S\zeta$ we shall mean that element obtained from ζ by performing the operation S. By $\sum_{i=1}^{g} a_i S_i \zeta$ we shall mean the unit $\prod_{i=1}^{g} (S_i\zeta)^{a_i}$. Hence the group ring with rational integral coefficients, which we shall denote by $[G]$, can be taken as a ring of operators [4] for the additive modules M/R' and U/R.

---

[1] C. G. Latimer, *American Journal of Mathematics*, vol. 56, no. 1 (1934), pp. 69-74.

[2] The author is indebted to the late Professor Emmy Noether for suggesting this method of treating the problem.

[3] H. Minkowski, *Goettinger Nachrichten* (1900), p. 90; H. Hasse, *Klassenkoerpertheorie, Anhang. Theorie der Einheiten.*

[4] E. Noether, *Mathematische Zeitschrift*, vol. 30, pp. 641-692.

We shall use a notation which will combine the cases Ω real and Ω imaginary. If in all that follows the sum $1/2\Sigma H$, the sum being over the elements of H, is replaced by the identity E, the imaginary case reduces to the real case. Let (G) denote the group ring with rational coefficients. The left ideal $(G)\Sigma H$ can be written as the direct sum of the two left ideals $(G)E_1$ and $(G)E_2$, where $E_1 = 1/g\Sigma G$ and $E_2 = 1/2\Sigma H - 1/g\Sigma G$ are idempotents. Write the group $G = H + S_2 H + \cdots + S_{r+1}H$. Using the integral group ring as a ring of operators, we shall set up an operator isomorphism between the $[G]$-left module $[G]mE_2 = [G]m/2\Sigma HE_2$, which we shall denote by $\mathfrak{M}$, and the additive module M/R'. Note that $1/2\Sigma H\eta_1 = \eta_1$, and label the conjugates of η_1 so that $S_i\, 1/2\Sigma H\eta_1 = \eta_i$, $i = 2, 3, \cdots, r+1$. Further denote the class of M/R' containing ζ by ζ^*. If we let

$$m/2\Sigma HE_2 = mE_2 \rightarrow \eta_1{}^*$$

and

$$m/2\, S_i\Sigma HE_2 = mS_iE_2 \rightarrow \eta_i{}^* \qquad (i = 2, 3, \cdots, r+1),$$

we have the desired operator isomorphism. Now

$$m/2 \sum_{i=1}^{r+1} S_i\Sigma HE_2 = 0 \rightarrow \sum_{i=1}^{r+1} \eta_i{}^* = 0$$

and

$$m/2 \sum_{i=1}^{r+1} a_iS_i\Sigma HE_2 \rightarrow \sum_{i=1}^{r+1} a_i\eta_i{}^*$$

equals zero if and only if the a_i are equal. Note that if

$$w_1 = m/2 \sum_{i=1}^{r+1} a_iS_i\Sigma HE_2 \rightarrow \zeta^*$$

and if

$$w_2 = m/2 \sum_{i=1}^{r+1} b_iS_i\Sigma HE_2 \rightarrow \zeta^*$$

then

$$w_1 - w_2 = m/2 \sum_{i=1}^{r+1} (a_i - b_i)S_i\Sigma HE_2 = 0,$$

and $a_i - b_i = k$, thus showing that w_1 and w_2 differ only by zero. We conclude that only r of the elements $m/2\Sigma HE_2$, $m/2S_2\Sigma HE_2, \cdots, m/2S_{r+1}\Sigma HE_2$ are linearly independent, and any r of them may be taken as an independent basis of $\mathfrak{M}$.

Since M is of finite index m in U, the m-th power of every unit in Ω is found in M. From the operator isomorphism we have established, we see that the image of every class of U/R is found in the $[G]$-left module $[G]E_2$.

If $\zeta_1, \zeta_2, \cdots, \zeta_r$ form a fundamental system of units in U, we have $u'_i \rightarrow m\xi_i^*$, u'_i an element of $\mathfrak{M}$, $i = 1, 2, \cdots, r$. Hence $u'_i/m = u_i \rightarrow \xi_i^*$. Thus all linear forms of the u_i with integral coefficients generate a module $\mathfrak{U}$ which is operator isomorphic to the additive module U/R, the integral group ring being the ring of operators. Note that if the elements $v_1, v_2, \cdots, v_r$ are elements of a second basis of $\mathfrak{U}$, then $(v_1, v_2, \cdots, v_r) = (u_1, u_2, \cdots, u_r)A$, where A is a matrix whose determinant equals ± 1, but then

$$(\zeta_1', \zeta_2', \cdots, \zeta_r') = (\zeta_1, \zeta_2, \cdots, \zeta_r)A,$$

and $\zeta_1', \zeta_2', \cdots, \zeta_r'$ form a fundamental system of units of U. Thus we have a one to one correspondence between the independent bases of $\mathfrak{U}$ and the fundamental systems of units in U.

The above operator isomorphism enables us to establish a criterion for the existence of a fundamental system of units composed of conjugates. If the modules $\mathfrak{M}$ and $\mathfrak{U}$ are operator isomorphic under the integral group ring, we see that a fundamental system of units composed of conjugates may be chosen. On the other hand if a fundamental system of units composed of conjugates exists, $\mathfrak{M}$ and $\mathfrak{U}$ are operator isomorphic. Hence we may state

THEOREM 1. *A necessary and sufficient condition for the existence of a fundamental system of units composed of conjugates is that the modules* $\mathfrak{M}$ *and* $\mathfrak{U}$ *be operator isomorphic under the integral group ring.*

Note that if Ω is real or if the subgroup H is normal, $\mathfrak{M}$ is a two sided $[G]$-module, generated by a single element. Hence in these cases our criterion for the existence of a fundamental system of units composed of conjugates is analogous to the one given by Latimer.[5]

We may also draw some conclusions concerning the existence of a fundamental system of real units. If Ω is imaginary and a fundamental system of real units exists, the conjugate imaginary $\bar{\eta}$ of a unit η equals $\rho^a\eta$, where ρ is a root of unity generating R. In this case every class of U/R is invariant under H. However, from the operator isomorphism we have set up, we see that every class of U/R is left invariant by H, if and only if H is invariant. Moreover, if every class of U/R is invariant under H, a unit and its conjugate imaginary are found in the same class. Hence we may seek fundamental systems of real units only in those fields whose maximum real subfield is normal.

THEOREM 2. *A necessary condition for the existence of a fundamental*

[5] C. G. Latimer, *loc. cit.*, Theorem 2.

system of real units in an imaginary field Ω normal with respect to the rational field is the invariance of the subgroup H of the Galois group G to which the maximum real subfield of Ω belongs. If H is invariant the conjugate imaginary of a unit η of Ω equals $\rho^a\eta$, where ρ is a root of unity.

Note that the conditions of the above theorem are satisfied in Abelian fields. *In all that follows we shall assume that H is an invariant subgroup of G.*

In case Ω is an imaginary cyclic field, the above theorem enables us to prove without using a matric representation of the group of units Latimer's theorem that a fundamental system of real units exists.[6] Let $\zeta_1, \zeta_2, \cdots, \zeta_r$ be a fundamental system of units. Since G is cyclic, H is invariant and $\bar{\zeta}_i = \rho^{a_i}\zeta_i$. Let $H = \{E, S^{r+1}\}$, S being a generating element of G. If $a_i = 2b_i$, the b_i integers, we choose the units $\rho^{b_i}\zeta_i$ as the elements of a fundamental system. These are real since $S^{r+1}(\rho^{b_i}\zeta_i) = \rho^{-b_i}\bar{\zeta}_i = \rho^{-b_i}\rho^{2b_i}\zeta_i = \rho^{b_i}\zeta_i$. We shall show that each a_i is even. Recall that R is generated by a primitive even root of unity. Let $S\rho = \rho^h$, then $S^{r+1}\rho = \rho^{h^{r+1}} = \rho^{-1}$ from which we conclude that h is odd. Now letting $S\zeta_i = \zeta_i^{(1)}$ and $S^v\zeta_i = \zeta_i^{(v)}$, we have

$$\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)}\zeta_i^{(r+1)} \cdots \zeta_i^{(2r+1)} = 1,$$

where $\zeta_i^{(r+j)}$ is the conjugate imaginary of $\zeta_i^{(j-1)}$ and $\zeta_i^{(0)} = \zeta_i$. Using the relation $\bar{\zeta}_i = \rho^{a_i}\zeta_i$, we obtain

$$(\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)})^2 = \rho^l$$

and since $\zeta_i \cdots \zeta_i^{(r)}$ is in Ω as well as its square

$$\zeta_i\zeta_i^{(1)} \cdots \zeta_i^{(r)} = \rho^k, \qquad k \text{ an integer.}$$

Operating by S on both sides of this last equation, we obtain

$$\zeta_i^{(1)}\zeta_i^{(2)} \cdots \zeta_i^{(r+1)} = \rho^{hk}$$

which gives

$$\zeta_i^{(r+1)} = \rho^{k(h-1)}\zeta_i.$$

Since h is odd we have proved our proposition.

We seek the image of those classes of U/R which contain units from a subfield Γ of Ω. Let the elements of K, the subgroup of index k to which the subfield Γ belongs, be $K_1 = E, K_2, \cdots, K_h$. Write G in terms of its v cosets with respect to the subgroup generated by K and H. If K con-

[6] C. G. Latimer, *loc. cit.*, Theorem 1.

tains H, $v = k$, while if K does not contain H, $v = k/2$. Let η^* be a class of U/R which contains a unit from the subfield Γ, and its image in $[G]E_2$ be

$$w = 1/2(\Sigma a_i K_i \Sigma H + \Sigma b_i K_i \Sigma H S_2 + \cdots + \Sigma v_i K_i \Sigma H S_v)E_2,$$

$i = 1, 2, \cdots, s$, where $s = h/2$ or h according as K does or does not contain H. Since $K\eta^* = \eta^*$, $Kw = w$, K any element of the group K. Equating coefficients in Kw and w, we have $a_i - a_k = t$, $b_i - b_k = t, \cdots, v_i - v_k = t$, the subscripts k being a permutation of the subscripts i. Hence

$$\sum_{i=1}^{s}(a_i - a_k) = st = 0, \; t = 0, \text{ and } a_i = a_k.$$

By operating in turn with each element of K, we see that the $a_i, b_i, \cdots, v_i$, respectively, are sets of equal integers. Thus to each class of U/R which contains a unit from the subfield Γ corresponds an element of the form:

$$1/2(a_1 \Sigma K_i \Sigma H + a_2 \Sigma K_i \Sigma H S_2 + \cdots + a_v \Sigma K_i \Sigma H S_v)E_2, \quad (i = 1, \cdots, s).$$

These elements are contained in a module $\mathfrak{K}$ which is a right-sided $[G]$-module. If we regard the v sums $1/2\sum_{i=1}^{s} K_i \Sigma H S_j E_2$ as elements, we see that every $v - 1$ of them form an independent basis for the module $\mathfrak{K}$. Note that $\mathfrak{K}$ may be generated by a single element $1/2\Sigma K_i \Sigma H E_2$, by multiplication on the right by $[G]$. Further if the subgroup K is invariant, $\mathfrak{K}$ is a two-sided $[G]$-module. We may summarize as follows:

THEOREM 3. *The images of those classes of U/R which contain units from the subfield belonging to the subgroup K of G are found in the right-sided $[G]$-module* $(1/2\sum_{i=1}^{s} K_i \Sigma H E_2)[G]$.

We now apply the theory developed thus far to the case where G is the direct product of two subgroups S and T of relatively prime orders s and t, respectively. Hence one subgroup contains H, say S does. We need to bear in mind that we are assuming H an invariant subgroup of G. For convenience let $s/2 = q$. Now $(G)E_2$ contains the following idempotents:

$I_1 = 1/s\Sigma S E_2 = 1/s\Sigma S - 1/g\Sigma G$,
$I_2 = 1/t\Sigma T E_2 = 1/(2t)\Sigma T \Sigma H - 1/g\Sigma G$, and
$I_3 = (1/2\Sigma H - 1/s\Sigma S - 1/t\Sigma T)E_2 = 1/2\Sigma H - 1/s\Sigma S - 1/(2t)\Sigma T\Sigma H + 1/g\Sigma G$,

such that $I_1 + I_2 + I_3 = E_2$, $I_i I_j = 0$, $i \neq j$, $I_i^2 = I_i$, $I_i E_2 = E_2 I_i = I_i$, $i = 1, 2, 3$. Hence $(G)E_2$ is the direct sum of $(G)I_1$, $(G)I_2$, and $(G)I_3$. Now consider the submodule $\mathfrak{W}$ formed by the direct sum of $\mathfrak{R} = [G]I_1$, $\mathfrak{S} = [G]I_2$, and $\mathfrak{T} = [G]I_3$. $\mathfrak{R}$ has an independent basis of $r_1 = t - 1$ of the t elements $I_1, T_2I_1, \cdots, T_tI_1$, $\mathfrak{S}$, an independent basis of $r_2 = q - 1$ of

the q elements $1/2\Sigma HI_2$, $1/2\Sigma HS_2I_2, \cdots, 1/2\Sigma HS_qI_2$, and $\mathfrak{T}$ an independent basis of $r_3 = r - t - q + 2$ elements. Since $\mathfrak{U}$ is a submodule of $\mathfrak{W}$, its basis can be expressed in the form:

$$
\begin{aligned}
u_i &= x_i & &(i = 1, 2, \cdots, r_1), \\
u_i &= x_i' + y_i & &(i = r_1 + 1, \cdots, r_1 + r_2), \\
u_i &= x_i'' + y_i' + z_i & &(i = r_1 + r_2 + 1, \cdots, r),
\end{aligned}
$$

where the x_i, x_i', x_i'' are linear combinations with integral coefficients of the base elements of $\mathfrak{R}$ and y_i and y_i' similar functions of the base elements of $\mathfrak{S}$, while the z_i are such functions of the base elements of $\mathfrak{T}$. Or, in other words, if we consider the matrix of the base elements of $\mathfrak{W}$, we may take zeros above the principal diagonal. Since the module $\mathfrak{U}$ admits $[G]$ as a ring of operators, we have,

$$\Sigma S u_i = s x_i'$$

and

$$t(E - 1/t\Sigma T)u_i = t x_i' \qquad (i = r_1 + 1, \cdots, r_1 + r_2).$$

Since the integers s and t are relatively prime, each of the x_i' is an element of $\mathfrak{U}$. Thus we may take a basis of $\mathfrak{U}$ in the above form with the elements x_i' replaced by zero.

As we have previously seen to an independent basis u_i of $\mathfrak{U}$ corresponds a fundamental system of units $\zeta_1, \zeta_2, \cdots, \zeta_r$ of Ω. Since the first r_1 of the u_i are invariant under the elements of the group S, while the next r_2 are invariant under the elements of T, the first r_1 of the ζ_i lie in the subfield Γ_1 which belongs to the subgroup S, and the next r_2 of the ζ_i lie in the subfield Γ_2 which belongs to the subgroup T. Since the ζ_i form a fundamental system of units in Ω, the first r_1 of the ζ_i form such a fundamental system in Γ_1 and the next r_2 form such a system in Γ_2. Hence we have

THEOREM 4. *Let Ω be an absolutely normal algebraic field in which the maximum real subfield is normal in case Ω is imaginary. If Ω is obtained by the composition of two subfields Γ_1 and Γ_2, of relatively prime degrees, Ω contains a fundamental system of units such that certain of them form a fundamental system in Γ_1 and certain others form such a system in Γ_2.*

If Γ_2 is again obtained by the composition of two fields of relatively prime degrees, we can apply the previous theorem to Γ_2 instead of Ω, using the quotient group G/T instead of G, and thus obtain a fundamental system of units of Γ_2 such that it contains fundamental systems from each of the two subfields which generate Γ_2. An Abelian group can always be written as the direct product of subgroups whose orders are powers of distinct primes. Hence the theorem may always be applied to Abelian fields.

BRYN MAWR COLLEGE.

A CALCULUS OF SEQUENCES.

By MORGAN WARD.

I. *Introduction.*

1. I propose in this paper to give a generalization of a large portion of the formal parts of algebraic analysis and the calculus of finite differences. The generalization consists in systematically replacing the ordinary binomial coefficient $n(n-1)\cdots(n-r+1)/1\cdot 2\cdots r$ by a "binomial coefficient to the base (u)," $[n,r]=u_n\cdot u_{n-1}\cdots u_{n-r+1}/u_1\cdot u_2\cdots u_r$ where (u): $u_0, u_1, u_2,\cdots$ is a fixed sequence of complex numbers subject to the restrictions $u_0=0$; $u_1=1$; $u_n\neq 0$, $n>1$. The exponential function for example is replaced by the formal series $1+\sum_{n=1}^{\infty} x^n/u_1\cdot u_2\cdots u_n$, differentiation by an operation which throws x^n into $[n,1]x^{n-1}$, and differencing by an operation which throws x^n into $\sum_{r=1}^{n}[n,r]x^{n-r}$.

The formula $n^r=(1+1+\cdots+1)^r=\sum_{(s)} r!/s_1!s_2!\cdots s_n!$ guides us in replacing the powers of rational integers where necessary by sums of multinomial coefficients to the base (u); for example, $\sum_{s=0}^{r}[n,r]$ may replace 2^r. We are thus enabled to generalize successfully a great variety of formulas involving the exponential functions, the Bernoulli numbers and polynomials.

2. In a series of papers which have appeared during the past thirty years [1],[1] F. H. Jackson has developed a somewhat similar extension of elementary analysis for the particular sequence (u) in which $u_n=(q^n-1)/(q-1)$ q a fixed complex number, $|q|\neq 1$. His results are based essentially on an identity of Euler's [2]:

$$(2.1)\qquad (x+y)(x+qy)\cdots(x+q^{n-1}y)=\sum_{r=0}^{n}[n,r]q^{r(r-1)/2}x^{n-r}y^r.$$

In effect he replaces the ordinary binomial coefficient by $[n,r]q^{r(r-1)/2}$. But the presence of this power of q introduces a lack of symmetry in his formulas,

[1] The numbers in square brackets refer to references at the end of the paper. Jackson wrote in all over thirty papers on this subject. We have listed only those directly connected with the present paper.

and leads to certain complications in defining the exponential functions.[2] He was not led furthermore to consider the developments of the calculus of finite differences which we give here. These appear to be among the most striking results of the entire theory.

The present work originated in an (unsuccessful) attempt to frame a definition of the Bernoulli numbers in Jackson's calculus to which the Staudt-von Klausen theorem might apply.

II. *Formal theory.*

3. Let

$$(u): \quad u_0 = 0, \quad u_1 = 1, \quad u_2, \cdots, u_n, \cdots$$

be a fixed sequence of complex numbers subject for the present to the single restriction $u_n \neq 0$, $n > 1$. For convenience, we shall write $[n]$ for u_n. We define:

$[n]!$ to be 1 if $n = 0$, and $[n][n-1]\cdots[1]$ if $n > 0$;
$[n, r]$ to be $[n]!/[r]![n-r]!$

where n, r are positive integers,[3] and $n \geqq r$. Then

$$[n, 0] = 1, \quad [n, 1] = [n], \quad [n, n-r] = [n, r].$$

We shall call $[n, r]$ a binomial coefficient to the base (u), or simply a *basic*[4] binomial coefficient.

We write $(x + y)^n$ for the polynomial $\sum_{r=0}^{n} [n, r] x^{n-r} y^r$. It is evident that

$$(x + y)^0 = 1, \quad (x + y)^1 = x + y, \quad (x + 0)^n = x^n$$
$$(cx + cy)^n = c^n (x + y)^n, \quad (x + y)^n = (y + x)^n.$$

From the identities[5]

$$(x - y)^{2n+1} = \sum_{r=0}^{n} (-1)^r [2n + 1, r] x^r y^r (x^{2n+1-2r} - y^{2n+1-r})$$
$$(x - y)^{2n} = \sum_{r=0}^{n-1} (-1)^r [2n, r] x^r y^r (x^{2n-2r} + y^{2n-2r}) + (-1)^n [2n, n] x^n y^n$$

---

[2] It is necessary to consider not only the series $1 + \sum_{1}^{\infty} x^n/u_1 u_2 \cdots u_n$ as an analogue of the exponential, but also the series $1 + \sum_{1}^{\infty} q^{n(n-1)/2} x^n/u_1 u_2 \cdots u_n$ with a corresponding complexity in the theory of the trigonometric functions.

[3] We count zero as a positive integer.

[4] This convenient terminology is due to F. H. Jackson.

[5] We write $(x - y)^n$ for $(x + (-y))^n = \sum_{r=0}^{n} [n, r] x^{n-r} (-y)^r$.

we see that

$$(x-x)^{2n+1}=0; \quad (n=0,1,2,\cdots).$$

On the other hand,

$$(x-x)^{2n}=x^{2n}(1-1)^{2n}=x^{2n}\{2\sum_{r=0}^{n-1}(-1)^r[2n,r]+(-1)^n[2n,n]\}$$

generally does *not* vanish. A sequence (u) such that

$$(1-1)^{2n}=0, \quad (n=1,2,3,\cdots)$$

will be said to be *normal.*

4. More generally, we define

$$(x_1+x_2+\cdots+x_k)^n \text{ to be } \sum_{(s)}\frac{[n]\,!}{[s_1]\,!\cdots[s_k]\,!}\,x_1^{s_1}\cdots x_k^{s_k}$$

where the summation is over all integers s satisfying the conditions

$$s_1+s_2+\cdots+s_k=n, \qquad 0\leqq s_i\leqq n.$$

If we denote this polynomial by $P_{kn}(x)=P_{kn}(x_1,x_2,\cdots,x_k)$ then it is a symmetric function of its k arguments, and if c is any constant, then

$$P_{kn}(cx_1,cx_2,\cdots,cx_k)=c^nP_{kn}(x_1,x_2,\cdots,x_k).$$

Furthermore,

(4.1) $$P_{k+1n}(x)=P_{kn}(x_1,x_2,\cdots,x_{k-1},x_k+x_{k+1}).$$

For consider $P_{2n}(x)=(x_1+x_2)^n$. We have

$$\begin{aligned}
(x_1+x_2)^n &= \sum_{t=0}^{n}\frac{[n]\,!}{[n-t]\,!\,[t]\,!}\,x_1^{n-t}x_2^{t}\\
(x_1+(x_2+x_3))^n &= \sum_{t=0}^{n}\frac{[n]\,!}{[n-t]\,!\,[t]\,!}\,x_1^{n-t}(x_2+x_3)^{t}\\
&= \sum_{t=0}^{n}\sum_{r=0}^{t}\frac{[n]\,!\,[t]\,!}{[n-t]\,!\,[t]\,!\,[t-r]\,!\,[r]\,!}\,x_1^{n-t}x_2^{t-r}x_3^{r}\\
&= \sum_{(s)}\frac{[n]\,!}{[s_1]\,!\,[s_2]\,!\,[s_3]\,!}\,x_1^{s_1}x_2^{s_2}x_3^{s_3},\\
&\qquad\qquad s_1+s_2+s_3=n,\ 0\leqq s_i\leqq n\\
&= (x_1+x_2+x_3)^n=P_{3n}(x).
\end{aligned}$$

Hence (4.1) is true for $k=2$. Its validity for any value of k follows by an easy induction.

It is evident that formula (4.1) can be extended so as to express $P_{k+ln}(x)$ in terms of $P_{kn}(x)$ in various ways. For example,

$$P_{4n}(x_1,x_2,x_3,x_4)=P_{2n}(x_1+x_2,x_3+x_4).$$

5. The numerical values of the polynomials $P_{kn}(x)$ when all of the arguments x_i are equal to plus one play an important rôle in the developments which are to follow. We shall write $\bar{k}^n$ for the number

$$P_{kn}(1, 1, \cdots, 1) = (1 + 1 + \cdots + 1)^n = \sum_{(s)} \frac{[n]!}{[s_1]! \cdots [s_k]!}.$$

We see from the formulas of section 4 that

$$\bar{3}^n = (\bar{2} + \bar{1})^n, \quad \bar{4}^n = (\bar{2} + \bar{2})^n = (\bar{3} + \bar{1})^n.$$

It is easily shown by induction that we have quite generally

$$\overline{r + s}^n = (\bar{r} + \bar{s})^n \tag{5.1}$$

where r and s are any positive integers.

If furthermore the sequence (u) is normal (section 3) then we can show by induction that (5.1) holds for any integral values of r and s. A somewhat longer induction establishes the formula

$$\overline{m_1 + m_2 + \cdots + m_t}^n = (\bar{m}_1 + \bar{m}_2 + \cdots + \bar{m}_t)^n \tag{5.2}$$

where if (u) is normal, $m_1, \cdots, m_t$ are any integers, but if (u) is not normal, the integers are to be positive.

In case (u) is normal, there is no gain in generality in replacing some of the plus signs in formula (5.2) by minus signs because we can show that

$$(\bar{r} - \bar{s})^n = (\bar{r} + \overline{-s})^n. \tag{5.3}$$

6. If $F(x)$ denotes the formal power series

$$F(x) = \sum_{n=0}^{\infty} c_n x^n, \tag{6.1}$$

we define $F(x + y)$ to mean the series

$$\sum_{n=0}^{\infty} c_n (x + y)^n = \sum_{n=0}^{\infty} \sum_{r=0}^{n} c_n [n, r] x^{n-r} y^r.$$

In like manner

$$\begin{aligned} F(x_1 + x_2 + \cdots + x_k) &= \sum_{n=0}^{\infty} c_n (x_1 + x_2 + \cdots + x_k)^n \\ &= \sum_{n=0}^{\infty} c_n P_{kn}(x). \end{aligned} \tag{6.2}$$

We have furthermore formal identities of the type

$$F(x_1 + x_2) = F(x_2 + x_1),$$
$$F(x_1 + x_2 + x_3) = F(x_1 + (x_2 + x_3))$$
$$F(x_1 + x_2 + x_3 + x_4) = F((x_1 + x_2) + (x_3 + x_4))$$

since the like identities hold for the polynomials $P_{kn}(x)$.

If in the series (6. 2) we make all the arguments x_i equal to x, the right side becomes $\sum_{n=0}^{\infty} c_n(x + x + \cdots + x)^n = \sum_{n=0}^{\infty} c_n\bar{k}^n x^n$. We shall accordingly denote the resulting series by $F(\bar{k}x)$. It is obvious then from the formula (5. 2) that

(6. 3) $$F(\overline{m_1 + m_2 + \cdots + m_t}x) = F(\overline{m}_1 x + \overline{m}_2 x + \cdots + \overline{m}_t x)$$

for suitably restricted integers m_i.

Let $F(x)$, $G(x)$, $H(x)$ be three formal power series in x. Then the following theorem is easily seen to be true.

THEOREM 6. 1. *If $F(x) = G(x) \pm H(x)$ and m, n are any positive integers, then $F(\overline{m}x) = G(\overline{m}x) \pm H(\overline{m}x)$ and*

$$F(\overline{m}x + \overline{n}y) = G(\overline{m}x + \overline{n}y) \pm H(\overline{m}x + \overline{n}y).$$

7. We next define an operator $D = D_x$ which transforms the formal power series (6. 1) into

(7. 1) $$F'(x) = DF(x) = \sum_{n=0}^{\infty} [n] c_n x^{n-1}.$$

In particular then, $Dx^n = [n]x^{n-1}$. The operator D is easily shown to be linear and distributive, and it converts a polynomial of degree n in x into one of degree $n - 1$.

If we define $F^{(r)}(x) = D^r F(x)$ recursively by $F^{(r+1)}(x) = DF^{(r)}(x)$; $F^{(0)}(x) = F(x)$, it easily follows that

$$\frac{F^{(r)}(x)}{[r]!} = \sum_{n=r}^{\infty} [n, r] c_n x^{n-r}.$$

The expansion $F(x + y) = \sum_{n=0}^{\infty} c_n (x + y)^n$ is formally replaceable by

(7. 2) $$F(x + y) = \sum_{n=0}^{\infty} \frac{F^{(n)}(x)}{[n]!} y^n.$$

We shall refer to (7. 2) as *Taylor's formula* for the base (u).

Finally, we note that

$$D_x{}^r F(x + y) = F^{(r)}(x + y).$$

8. As a simple concrete example of such an operator D, let us assume that the sequence (u) is a linear recurring series of order k whose associated polynomial $x^k - a_1x^{k-1} - \cdots - a_k$ has k distinct roots $\alpha_1, \alpha_2, \cdots, \alpha_k$. Then u_n is of the form

$$u_n = \beta_1\alpha_1{}^n + \beta_2\alpha_2{}^n + \cdots + \beta_k\alpha_k{}^n$$

where the constants α and β are subject to the conditions

$$\beta_1 + \beta_2 + \cdots + \beta_k = 0, \quad \beta_1\alpha_1 + \beta_2\alpha_2 + \cdots + \beta_k\alpha_k = 1, \quad u_n \neq 0, n > 1.$$

It is obvious then that

$$DF(x) = (\beta_1F(\alpha_1x) + \cdots + \beta_kF(\alpha_kx))/(\beta_1\alpha_1x + \cdots + \beta_k\alpha_kx).$$

This operator can therefore be applied to any function of x regular at $x = 0$, and transforms it into another function regular at $x = 0$.

In particular, if $k = 2$, $\alpha_1 = q$, $\alpha_2 = 1$, $\beta_1 = (q-1)^{-1}$, $\beta_2 = -\beta_1$, where q is not a root of unity,

$$DF(x) = \frac{F(qx) - F(x)}{qx - x}$$

is the operation of q-differencing.

In case some of the roots α of the polynomial associated with the recurrence relation are repeated, a similar but more complicated formula for $DF(x)$ may be given which involves both $F(x)$ and its ordinary derivatives. For example, if $k = 2$ and $\alpha_1 = \alpha_2 \neq 0$, $u_n = n\alpha_1{}^{n-1}$ and $DF(x) = \dfrac{1}{\alpha_1}\dfrac{dF(\alpha_1x)}{dx}$.

III. *The exponential and trigonometric functions.*

9. We shall now assume that the sequence (u) is chosen in such a manner that the series

(9.1) $$\mathcal{E}(x) = \sum_{n=0}^{\infty} \frac{x^n}{[n]!}$$

is convergent in the neighborhood of $x = 0$. It accordingly is an element of an analytic function of x which we shall call the *basic exponential.* There exists then a positive number ρ such that the series (9.1) converges absolutely within the circle $|x| = \rho$.

The basic exponential has the following properties for sufficiently small absolute values of its arguments x, y, x_i:

(9.2) $$D\mathcal{E}(x) = \mathcal{E}(x), \quad \mathcal{E}^{(n)}(cx) = c^n\mathcal{E}(cx), c \quad \text{a constant,}$$

(9.21) $$\mathcal{E}(x+y) = \mathcal{E}(x)\mathcal{E}(y),$$

(9.22) $$\mathcal{E}(x_1 + x_2 + \cdots + x_k) = \mathcal{E}(x_1)\mathcal{E}(x_2)\cdots\mathcal{E}(x_k).$$

Consider for example the formula (9.21). That it is formally true is immediately obvious from the basic Taylor's formula (7.2). For

$$\mathcal{E}(x+y) = \sum_{n=0}^{\infty} \frac{\mathcal{E}^{(n)}(x) y^n}{[n]!} = \mathcal{E}(x)\mathcal{E}(y)$$

since by (9.2), $\mathcal{E}^{(n)}(x) = \mathcal{E}(x)$.

But the series

$$\sum_{n=0}^{\infty} \frac{(x+y)^n}{[n]!} = \sum_{n=0}^{\infty} \sum_{r=0}^{n} \frac{x^{n-r} y^r}{[n-r]!\,[r]!}$$

is in fact the Cauchy product of the series $\sum \frac{x^n}{[n]!}, \sum \frac{y^n}{[n]!}$ so that the formula is actually true provided the latter two series are both absolutely convergent. And by our initial hypothesis, both series converge absolutely if $|x| < \rho$, $|y| < \rho$.

10. The trigonometric and hyperbolic functions are defined by Euler's formulas:

$$\text{(10.1)} \quad \begin{aligned} &\sin(x) = \frac{\mathcal{E}(ix) - \mathcal{E}(-ix)}{2i}, \quad \cos(x) = \frac{\mathcal{E}(ix) + \mathcal{E}(-ix)}{2}, \\ &\sinh(x) = -i\sin(ix), \quad \cosh(x) = \cos(ix). \end{aligned}$$

Among the many formal analogies with the ordinary trigonometric functions, we shall merely note here:

$$\begin{aligned} \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y), \\ \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y), \\ D\sin(x) &= \cos(x), \qquad D\cos(x) = -\sin(x). \end{aligned}$$

As a consequence of the last two formulas, we see that both $\sin(x)$ and $\cos(x)$ satisfy the basic differential equation $y^{(2)}(x) + y(x) = 0$.

On the other hand

$$\text{(10.2)} \qquad \sin^2(x) + \cos^2(x) = \mathcal{E}(ix)\mathcal{E}(-ix)$$

and in general, $\mathcal{E}(ix)\mathcal{E}(-ix) \neq 1$.

The remaining trigonometric and hyperbolic functions are defined in terms of the basic sine and cosine as in the ordinary case.

11. It is possible to give analogues of De Moivre's and Simpson's formulas. For in formula (9.22), take $k = n$ and let $x_1 = x_2 = \cdots = x_n = i\theta$. Then with the notation explained in section 6,

$$\text{(11.1)} \qquad \mathcal{E}(\bar{n}i\theta) = (\mathcal{E}(i\theta))^n.$$

Hence we obtain from the formulas (10. 1) and theorem 6. 1 the basic form of De Moivre's formula,[6] $\cos(\bar{n}\theta) + i\sin(\bar{n}\theta) = (\cos(\theta) + i\sin(\theta))^n$.

THEOREM 11. 1. *The sequence* (u) *is normal when and only when* $\mathcal{E}(x)\mathcal{E}(-x) = 1$ *or when and only when* $\sin^2(x) + \cos^2(x) = 1$.

For by formula (9. 21), and the previous definitions, if $|x| < \rho$,

$$\mathcal{E}(x)\mathcal{E}(-x) = \mathcal{E}(x - x) = \sum_{n=0}^{\infty} \frac{(x-x)^n}{[n]!} = \sum_{n=0}^{\infty} \frac{(1-1)^n}{[n]!} x^n.$$

Now we have seen in section 3 that $(1-1)^{2n+1} = 0$, $(n = 0, 1, 2, \cdots)$. Therefore

$$\mathcal{E}(x)\mathcal{E}(-x) = 1 + \sum_{n=0}^{\infty} \frac{(1-1)^{2n}}{[2n]!} x^{2n}.$$

Hence the first part of the theorem follows. The second part of the theorem is an immediate consequence of formula (10. 2).

Let us assume now that (u) is normal. We see from formula (11. 1) that

$$\begin{aligned} \mathcal{E}(\overline{n+2}i\theta) &= \mathcal{E}(i\theta)\mathcal{E}(\overline{n+1}i\theta), \\ \mathcal{E}(i\theta)\mathcal{E}(\bar{n}i\theta) &= \mathcal{E}(\overline{n+1}i\theta). \end{aligned} \tag{11. 3}$$

But by theorem 11. 1, $\mathcal{E}(-i\theta)\mathcal{E}(i\theta) = 1$. Therefore this last equation may be written

$$\mathcal{E}(\bar{n}i\theta) = \mathcal{E}(-i\theta)\mathcal{E}(\overline{n+1}i\theta). \tag{11. 31}$$

On adding and subtracting the two formulas (11. 3), (11. 31) and applying (10. 1) and theorem (6. 1), we obtain the basic Simpson's formulas:

$$\begin{aligned} \cos(\overline{n+2}\theta) &= 2\cos(\theta)\cos(\overline{n+1}\theta) - \cos(\bar{n}\theta), \\ \sin(\overline{n+2}\theta) &= 2\cos(\theta)\sin(\overline{n+1}\theta) - \sin(\bar{n}\theta). \end{aligned}$$

12. In order that the results of the previous section may have more than a purely formal significance, it is necessary to show that we can choose the sequence (u) so that (u) is normal and so that $E(\bar{n}x)$ is an entire function of x for any integer n. Since $E(\overline{-n}x) = E(-\bar{n}x)$, $E(\bar{n}x) = (E(\bar{1}x))^n$, $E(\bar{1}x) = E(x)$, we need only consider the case when $n = +1$.

Now it is easy to show that the most general solution of the functional equation

$$\Phi(x)\Phi(-x) = 1 \tag{12. 1}$$

[6] It should be noted here that $(\cos(\theta) + i\sin(\theta))^n$ stands for the product $(\cos(\theta) + i\sin(\theta))(\cos(\theta) + i\sin(\theta))\cdots$ taken to n factors, and not for the result of substituting $\cos(\theta)$ for x_1, and $i\sin(\theta)$ for x_2 in the polynomial $P_{2n}(x) = (x_1 + x_2)^n$.

which is regular at the origin is of the form

$$\Phi(x) = \pm \exp.(x\Psi(x^2)) \tag{12.2}$$

where $\Psi(x)$ is regular at the origin. But by theorem 11. 1, (u) is normal when and only when $\mathcal{E}(x)$ is a solution of (12. 1). Since $\mathcal{E}(x)$ was assumed to be regular at the origin, $\mathcal{E}(x)$ must be of the form (12. 2) where $\Psi(x)$ is an entire function of x. We must also satisfy the conditions [7]

$$\mathcal{E}(0) = 1, \quad \mathcal{E}'(0) = 1, \quad \frac{\mathcal{E}^{(n)}(0)}{n!} = \frac{1}{u_1 u_2 \cdots u_n} \neq 0,$$

as then $u_n = n\mathcal{E}^{(n-1)}(0)/\mathcal{E}^{(n)}(0) \neq 0 \quad (n = 1, 2, \cdots)$ and $u_1 = 1$.

It will therefore suffice to choose for $\Psi(x)$ an entire function $G(x)$ with a series expansion of the form $G(x) = 1 + \sum_{n=1}^{\infty} g_n x^n$ where the quantities g_n are all real and non-negative. The ordinary case ensues on taking all the quantities g_n equal to zero.

13. If we assume that $\mathcal{E}(x)$ is an entire function satisfying the condition $\mathcal{E}(x)\mathcal{E}(-x) = 1$, we can generalize the periodic properties of the exponential function. For since $\mathcal{E}(x)$ never vanishes, by Picard's theorem there exists a complex number $\lambda \neq 0$ such that $\mathcal{E}(\lambda) = 1$. But then if n is a positive integer,

$$\begin{aligned} \mathcal{E}(x + \overline{n}\lambda) &= \mathcal{E}(x)\mathcal{E}(\overline{n}\lambda) = \mathcal{E}(x)(\mathcal{E}(\lambda))^n = \mathcal{E}(x), \\ \mathcal{E}(x) &= \mathcal{E}(x - \overline{n}\lambda + \overline{n}\lambda) = \mathcal{E}(x - \overline{n}\lambda)\mathcal{E}(\overline{n}\lambda) = \mathcal{E}(x - \overline{n}\lambda). \end{aligned}$$

We have therefore proved the following theorem.

THEOREM 13. 1. *If (u) is a normal sequence so chosen that the basic exponential function $\mathcal{E}(x)$ is an entire function of x, and if $\lambda \neq 0$ is any zero of the function $\mathcal{E}(x) - 1$, and m any integer, then*

$$\mathcal{E}(x + \overline{m}\lambda) = \mathcal{E}(x).$$

Furthermore one such zero λ always exists.

On utilizing the formulas of section 10, we can easily show that under the hypotheses of theorem 13. 1, we also have

$$\sin(x + \overline{m}i\lambda) = \sin(x), \quad \cos(x + \overline{m}i\lambda) = \cos(x).$$

[7] The superscripts here denote ordinary differentiation.

IV. *The calculus of finite differences.*

14. Let (u) now be subject only to the conditions $u_0 = 0$, $u_1 = 1$, $u_n \neq 0$, $n \neq 0$. We shall denote by $\mathfrak{R}$ the ring of all polynomials in x with coefficients in the field of all complex numbers.

If

$$\phi = \phi(x) = \sum_{r=0}^{n} a_{n-r} x^r, \qquad a_0 \neq 0 \tag{14.1}$$

is any element of $\mathfrak{R}$ of degree n, we define the basic displacement symbol E by

$$\begin{aligned} E\phi(x) &= \phi(x+1) = \sum_{r=0}^{n} \sum_{s=0}^{r} a_{n-r}[r, s] x^{r-s} \\ E^{t+1}\phi(x) &= E(E^t\phi(x)), \quad E^0\phi(x) = \phi(x) \end{aligned} \tag{14.2}$$

where t is any positive integer.

It is obvious that E is a linear and distributive operator over $\mathfrak{R}$, and it may readily be shown that

$$E^t\phi(x) = \phi(x + \bar{t}). \tag{14.3}$$

If (u) is normal, formula (14.3) holds for all integral values of t.

15. The *basic difference operator* Δ is defined to be $E - 1$, where 1 stands for the identity operator over $\mathfrak{R}$. The following properties of Δ may be mentioned.

(i) Δ is linear and distributive over $\mathfrak{R}$, and converts an element of $\mathfrak{R}$ of degree n into one of degree $n - 1$. Moreover E, Δ and D are commutative over $\mathfrak{R}$.

(ii) The only solutions of $\Delta\phi = 0$ lying in $\mathfrak{R}$ are $\phi =$ a constant.

(iii) $\Delta^t\phi(x) = \sum_{s=0}^{t} (-1)^s \binom{t}{s} \phi(x + \bar{s})$.

(iv) We have the operational identity over $\mathfrak{R}$

$$\Delta = \mathcal{E}(D) - 1 \tag{15.1}$$

where formally $$\mathcal{E}(D) = \sum_{n=0}^{\infty} \frac{D^n}{[n]\,!}.$$

The last one of these properties is the only one requiring comment. If ϕ of formula (14.1) is operated on by D of section 7, then

$$\frac{D^s\phi(x)}{[s]\,!} = 0, \quad s > n; \qquad = \sum_{r=s}^{n} [r, s] a_{n-r} x^{r-s}, \quad s \leqq n.$$

Hence

$$\mathcal{E}(D)\phi(x) = \sum_{s=0}^{n} \sum_{r=s}^{n} [r, s] a_{n-r} x^{r-s} = \sum_{r=0}^{n} \sum_{s=0}^{r} [r, s] a_{n-r} x^{r-s} = E\phi(x)$$

by formula (14. 2), so that (15. 1) follows.

16. The *basic Bernouilli numbers* $B_0, B_1, \cdots, B_n, \cdots$ are defined by the recurrences

$$B_0 = 1; \quad (B+1)^n - B^n = 0,\ n > 1; \quad (B+1)^1 - B^1 = 1.$$

Here after expansion the exponents of B are to be degraded into suffices as in the usual theory [3].

The *basic Bernoulli polynomials* $B_n(z)$ may then be defined by

$$B_n(z) = (z+B)^n, \qquad (n = 0, 1, \cdots)$$

or non-symbolically,

$$B_n(z) = \sum_{r=0}^{n} [n, r] B_r z^{n-r}.$$

The following results [4] may be established precisely as in the ordinary theory.

(16. 1) $\quad B_n(0) = B_n, \quad B_n(1) = B_n, \quad n \neq 1; \quad B_1(1) = B_1 + 1.$

(16. 2) $\quad B_n(x+y) = \sum_{r=0}^{n} [n, r] x^r B_{n-r}(y).$

THEOREM 16. 1. *If $\phi'(x) = D\phi(x)$ denotes the basic derivative of the polynomial $\phi(x)$, then a polynomial solution of the difference equation*

$$\Delta\Psi(x) = \phi'(x)$$

is given by

(16. 3) $$\Psi(x) = \phi(x+B).$$

(16. 31) $$\phi(x+B) = \sum_{r=0}^{n} \frac{\phi^{(r)}(x)}{[r]!} B_r = \sum_{r=0}^{n} \frac{\phi^{(r)}(0)}{[r]!} B_r(x).$$

(16. 32) $$\Delta B_n(x) = [n] x^{n-1}.$$

THEOREM 16. 2. *If the sequence (u) be chosen so that the series (9. 1) for $\mathcal{E}(x)$ is convergent near $x = 0$ then for sufficiently small values of $|t|$ and $|x|$*

(16. 4) $$\sum_{n=0}^{\infty} \frac{B_n t^n}{[n]!} = \frac{t}{\mathcal{E}(t) - 1}, \quad \sum_{n=0}^{\infty} \frac{B_n(x) t^n}{[n]!} = \frac{t\mathcal{E}(xt)}{\mathcal{E}(t) - 1}.$$

(16. 5) $$\overline{1}^r + \overline{2}^r + \cdots + \overline{n-1}^r = \frac{B_{r+1}(\overline{n}) - B_{r+1}}{[r+1]},$$

if r is a positive integer $\geqq 1$.

To prove the last written formula for example, we observe by theorem 6.1 that (16.32) implies that $B_{r+1}(\overline{s+1}) - B_{r+1}(\bar{s}) = [r+1]\bar{s}^r$, s a positive integer. On summing this equation with respect to s from 0 to $n-1$, we obtain (16.5).

THEOREM 16.3. $B_{2n} = 0$ $(n = 1, 2, 3, \cdots)$ *when and only when*

$$B_n(1-z) = (-1)^n B_n(z), \quad (n = 2, 3, \cdots).$$

If moreover the series (9.1) *for* $\mathcal{E}(x)$ *converges for some* $x \neq 0$, *then*

$$B_{2n} = 0, \quad n \leqq 1$$

when and only when (u) *is normal.*

The equivalences stated follow immediately from formulas (16.1), (16.2) and (16.4).

We plan to give elsewhere a detailed treatment of the basic analogues for the numbers of Euler, Genocchi, Lucas and Stirling and their associated polynomials and difference operators.

REFERENCES.

1. F. H. Jackson, *American Journal of Mathematics*, vol. 32 (1910), pp. 305-314; *Messenger of Mathematics*, vol. 39 (1910), pp. 26-28, vol. 38 (1909), pp. 57-61, 62-64; *Proceedings of the Edinburg Mathematical Society*, vol. 22 (1904), pp. 28-39.

2. L. Euler, *Introductio in Analysin Infinitorum* (1748), chapter VII; Netto, *Combinatorik*, 2d. ed. (1927), p. 143.

3. D. H. Lehmer, *Annals of Mathematics* (2), vol. 36, no. 3, July (1935), p. 639 and references on p. 637.

4. Norlund, *Differenzenrechnung*, Berlin (1924), chapter II.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

ON THE TRANSFORMATION AND CONVERGENCE OF CONTINUED FRACTIONS.

By Walter Leighton and H. S. Wall.

1. *Introduction.* Convergence criteria for the continued fraction

$$y_0 + 1/y_1 + 1/y_2 + \cdots, \tag{1}$$

in which the y_n are complex numbers have been the object of study of many writers. E. B. Van Vleck [1] proved that if $y_1, y_3, y_5, \cdots$ are not all zero and if there exists a number $\epsilon > 0$ such that

$$-\pi/2 + \epsilon \leqq \arg y_n \leqq \pi/2 - \epsilon, \qquad (n = 1, 2, 3, \cdots)$$

a necessary and sufficient condition for the convergence of the continued fraction (1) is that the series $\Sigma \mid y_n \mid$ diverge. Pringsheim [2] showed that the conditions

$$\mid 1/y_{n-1}y_n \mid \leqq 1/4, \qquad (n = 2, 3, 4, \cdots) \tag{2}$$

are sufficient for the convergence of (1), with $\arg y_n$ unrestricted. Szász [3] pointed out that 1/4 is the largest number that can be used in (2).

When $\arg y_n$ is not restricted and the numbers $1/(y_n y_{n-1})$ do not satisfy the inequalities (2), the question of convergence of the continued fraction remains in a large measure unanswered.

In § 3 of this paper we have approached this problem by considering a new and very general transformation of a continued fraction. This approach has led to convergence theorems which are new and of a character fundamentally different from previously existing theorems. Some of these results are:

The continued fraction (1) *converges if*

$$0 < \mid 1/y_{2n}y_{2n+1} \mid \leqq 1/4, \quad \mid 1/y_{2n-1}y_{2n} \mid \geqq 25/4, \quad (n = 1, 2, 3, \cdots); \tag{3}$$

or if

$$G_n \geqq \mid 1/y_{2n}y_{2n+1} \mid \geqq 25/4, \quad \mid 1/y_{2n-1}y_{2n} \mid \geqq 4(1 + G_n)(1 + G_{n-1}), \quad (n = 1, 2, 3, \cdots), \tag{4}$$

[1] E. B. Van Vleck, "On the convergence of continued fractions with complex elements," *Transactions of the American Mathematical Society*, vol. 2 (1901).

[2] O. Perron, *Die Lehre von den Kettenbrüchen*, 1 st. edition, p. 254.

[3] O. Szász, "Über die Erhaltung der Konvergenz Kettenbrüche bei independenter Veränderlichkeit aller ihrer Elemente," *Crelle*, Bd. 147 (1916), pp. 132-160.

where $G_0 = 0$ *and* $G_1, G_2, G_3, \cdots$ *is any sequence of positive numbers.*

The following theorem is of a different type.

Let $h_1, h_2, h_3, \cdots$ *be any constants not zero such that for some integer* $k \geq 0$

$$\lim_{n=\infty} h_{2n}h_{2n+2k-1} = 0, \qquad \lim_{n=\infty} h_{2n}h_{2n+2k+1} = \infty.$$

Then corresponding to every finite region R *in the complex* z*-plane at a positive distance from the origin, there exists a positive integer* N *such that if* $n \geq N$, *the continued fraction*

$$h_{2n} + z/h_{2n+1} + z/h_{2n+2} + \cdots$$

converges uniformly over R.

Still another kind of theorem is the following.

Let $q_1, q_2, q_3, \cdots$ *be constants not zero. Let* B *represent an increasing sequence of indices* $n_1 < n_2 < n_3 < \cdots$ *in which* $n_i - n_{i-1} \geq 2$, $i = 2, 3, 4, \cdots$, *and let* A *represent the sequence of indices not in* B. *If* B *is finite and* $\lim q_n = 0$, *or if*

$$\lim_{B} q_n = \infty, \qquad \lim_{A} q_n = 0,$$

the continued fraction

$$1 + q_n z/1 + q_{n+1}z/1 + \cdots$$

converges uniformly over a region R *of the type described above, if* n *is taken sufficiently large.*

2. *An important example.* In this section we indicate the complexity of the general convergence problem by exhibiting a convergent continued fraction (1) for which the numbers

$$x_n = 1/y_{n-1}y_n$$

are everywhere dense in the complex plane.

Let $a_1, a_2, a_3, \cdots$ be an infinite sequence of complex numbers not zero and set

$$A_n = a_1 + a_2 + \cdots + a_n.$$

Form the continued fraction (1) in which

$$(5) \quad y_1 = \frac{2}{a_1}, \quad y_{2n} = -a_n, \; y_{2n+1} = \frac{2(a_n + a_{n+1})}{a_n a_{n+1}}, \quad (n = 1, 2, 3, \cdots).$$

Its $2n$-th and $(2n+1)$-th convergents are the quantities A_n and $A_n+(a_{n+1}/2)$, respectively. Hence the continued fraction converges if and only if the series Σa_n converges.

Now let S denote any countable set of complex numbers $s_1, s_2, s_3, \cdots$ not containing 0 or $-1/2$. It is clear that given any positive number M and a number r such that $0<r<1$, one can find a sequence of numbers $k_1, k_2, k_3, \cdots$ for which the conditions

$$\left|\frac{2s_n}{k_n}\right|<M, \quad \left|\frac{1+2s_n}{k_n}\right|<M, \quad k_n r^2+2k_n r^2 s_{n+1}-2s_n k_{n+1}\neq 0, \qquad (n=1,2,3,\cdots),$$

are satisfied.

If we take

$$a_{2n}=\frac{-2s_n r^{2n}}{k_n}, \quad a_{2n-1}=\frac{(1+2s_n)r^{2n}}{k_n}, \qquad (n=1,2,3,\cdots)$$

the series Σa_n and the continued fraction (1), with the numbers y_n given by (5), will converge. Moreover,

$$x_{4n+3}=s_{n+1}, \qquad (n=0,1,2,\cdots).$$

Thus *the set S which may be taken everywhere dense in the complex plane, is contained in the set of elements $x_1, x_2, x_3, \cdots$ of a convergent continued fraction.*

From this example one concludes that *the convergence of the continued fraction* (1) *does not in general depend upon the magnitude of the* $|x_n|$, *but upon the relative magnitude of the numbers in the ordered sequence*

$$|x_2|, \quad |x_3|, \quad |x_4|, \cdots.$$

This principle would lead one to expect to find criteria such as those mentioned at the end of § 1, in which precisely this "relative magnitude" plays the dominant rôle.

3. *A transformation of continued fractions, and an extension of the Pringsheim criteria.* Let

(6) $$\xi\equiv(x_1, x_2, \cdots; y_0, y_1, y_2, \cdots)=y_0+x_1/y_1+x_2/y_2+\cdots \qquad (x_n\neq 0)$$

and

(7) $$\eta\equiv(a_1, a_2, \cdots; b_0, b_1, b_2, \cdots)=b_0+a_1/b_1+a_2/b_2+\cdots \qquad (a_n\neq 0)$$

be any two infinite continued fractions. Denote their n-th convergents by X_n/Y_n and A_n/B_n, respectively, where the quantities X_n, Y_n, A_n, B_n are given by the usual recursion relations. As is well-known,

$$\begin{vmatrix} X_n & X_{n+1} \\ Y_n & Y_{n+1} \end{vmatrix} = (-1)^{n+1}x_1x_2\cdots x_{n+1},$$

and hence, one can always determine numbers α_n, β_n such that

$$\begin{aligned} A_n &= \alpha_n X_n + \beta_n X_{n+1}, \\ B_n &= \alpha_n Y_n + \beta_n Y_{n+1}, \end{aligned} \qquad (n = 0, 1, 2, \cdots). \tag{8}$$

Conversely, corresponding to any continued fraction (6), let numbers α_n, β_n be any set for which

$$D_n \equiv \alpha_{n-1}(\alpha_n + \beta_n y_{n+1}) - \beta_n\beta_{n-1}x_{n+1} \neq 0,\ n = 0, 1, 2, \cdots;\ \alpha_1 = 1, \beta_{-1} = 0.$$

If then we put

$$E_n \equiv \alpha_n[\alpha_{n-2}y_n - \beta_{n-2}x_n] + \beta_n[\alpha_{n-2}(y_n y_{n+1} + x_{n+1}) - \beta_{n-2}x_n y_{n+1}],$$
$$n = 1, 2, 3, \cdots;\ \alpha_{-1} = 1, \beta_{-1} = 0,$$

the continued fraction (7) with

$$\begin{aligned} &b_0 = y_0 + \beta_0 x_1/D_0,\ a_1 = x_1 D_1/D_0^2, \\ &a_n = x_n D_n/D_{n-1}, && (n = 2, 3, 4, \cdots), \\ &b_n = E_n/D_{n-1}, && (n = 1, 2, 3, \cdots), \end{aligned} \tag{9}$$

will have the property that

$$\begin{aligned} D_0 A_n &= \alpha_n X_n + \beta_n X_{n+1}, \\ D_0 B_n &= \alpha_n Y_n + \beta_n Y_{n+1}, \end{aligned} \qquad (n = 0, 1, 2, \cdots). \tag{10}$$

Equations (10) are readily established by mathematical induction.

We shall consider the equations (9) as constituting a transformation T of ξ into η, and write

$$(a_1, a_2, \cdots;\ b_0, b, \cdots) = T(x_1, x_2, \cdots;\ y_0, y_1, \cdots)$$

or

$$\eta = T\xi.$$

The transformation is determined by the α_n, β_n, so that we shall write $T = T[\alpha_n; \beta_n]$. This notation will simplify the later discussion.

The use of special cases of this transformation such as the so-called "equivalent transformation" where

$$\alpha_0 = 1,\ \alpha_n = c_1c_2\cdots c_n \qquad (c_i \neq 0);\ \beta_n = 0,$$

is well-known.[4] We shall use the more general transformation in a comparable manner.

In order to apply the transformation T for obtaining convergence criteria for continued fractions it is important to establish conditions under which the convergence of $\eta = T\xi$ implies the convergence of ξ. For that purpose we employ the identity

$$(11) \qquad \frac{A_n}{B_n} - \frac{X_n}{Y_n} = \left(\frac{A_{n-1}}{B_{n-1}} - \frac{A_n}{B_n}\right)\left(\frac{\alpha_{n-1}B_n}{\beta_n x_{n+1} B_{n-1}} - 1\right)^{-1}.$$

We have the following result.

THEOREM 1. *If $\eta = T\xi$ converges, ξ converges to the same limit as η, provided, there exists a positive integer N and a positive constant c such that*

$$| \alpha_{n-1}B_n/\beta_n x_{n+1}B_{n-1} - 1 | \geqq C,$$

if $n > N$.

The proof is an easy consequence of (11).

If the elements a_n, b_n of η satisfy the inequalities

$$(12) \qquad | b_n | \geqq | a_n | + 1, \qquad (n = 1, 2, 3, \cdots)$$

it may be shown[5] that η converges and the value of $\eta - b_0$ is numerically ≤ 1, and is < 1 if the inequality sign holds at least once in (12). Moreover, $| B_n/B_{n-1} | \geq 1$; hence, on referring to Theorem 1 we have the following theorem.

THEOREM 2. *Let $T[\alpha_n; \beta_n]$ be a transformation which carries the infinite continued fraction $\xi = (x_1, x_2, \cdots; y_0, y_1, y_2, \cdots)$, $(x_n \neq 0)$, into another infinite continued fraction $\eta = (a_1, a_2, \cdots; b_0, b_1, b_2, \cdots)$, $(a_n \neq 0)$, whose elements a_n, b_n satisfy* (12). *If there is a positive integer N and a number $s > 1$ such that for $\alpha_{n-1}\beta_n \neq 0$ and $n > N$*

$$(13) \qquad | \alpha_{n-1}/\beta_n x_{n+1} | \geqq s,$$

then ξ and η converge to a common limit. The numerical value of $\xi - b_0$ is ≤ 1 and is < 1 if the inequality sign holds at least once in (12).

The foregoing theorem may be extended as follows. Let $c_1, c_2, c_3, \cdots$ be a sequence of constants none of which is zero. The continued fraction[6] $\eta' = (c_1 a_1, c_1 c_2 a_2, c_2 c_3 a_3, \cdots; b_0, c_1 b_1, c_2 b_2, \cdots)$ converges if and only if η

[4] Perron, *loc. cit.*, p. 196.

[5] Perron, *ibid.*, pp. 254-255.

[6] Perron, *ibid.*, p. 196.

converges, and to the same limit. Let A'_n/B'_n be the n-th convergent of $(c_1c_2a_2, c_2c_3a_3, \cdots; 0, c_2b_2, c_3b_3, \cdots)$. If the latter converges to a value different from $-c_1b_1$, η' converges. If then we assume that

$$(14) \qquad |c_1b_1| \geqq 1, \quad |c_nb_n| \geqq |c_{n-1}c_na_n| + 1, \qquad (n = 2, 3, 4, \cdots),$$

where inequality holds at least once, it must follow that the limit

$$(15) \qquad \lim_{n=\infty} [c_1b_1 + A'_n/B'_n]$$

exists and is $\neq 0$. If we set

$$v_n = c_1b_1 + A'_n/B'_n$$

and denote the limit (15) by v, we observe that

$$\eta = b_0 + (c_1a_1/v),$$
$$\left| \frac{\alpha_{n-1}B_n}{\beta_n x_{n+1} B_{n-1}} \right| = \left| \frac{\alpha_{n-1}B'_{n-1}v_{n-1}}{c_n\beta_n x_{n+1} B'_{n-2}v_{n-2}} \right|,$$

and that $|B'_{n-1}/B'_{n-2}| \geq 1$. We have thus proved the following theorem.

THEOREM 3. *Let $T[\alpha_n; \beta_n]$ be a transformation which carries the infinite continued fraction $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$, $(x_n \neq 0)$, into another infinite continued fraction $\eta = (a_1, a_2, \cdots; b_0, b_1, \cdots)$ whose elements satisfy conditions* (14) *with actual inequality holding at least once for some set of numbers $c_1, c_2, c_3, \cdots$, $(c_i \neq 0)$. If there exists a positive integer N and a number $s > 1$ such that*

$$(16) \qquad \left| \frac{\alpha_{n+1}}{c_n\beta_n x_{n+1}} \right| \geqq s$$

for all values of $n > N$ for which $\alpha_{n-1}\beta_n \neq 0$, then ξ and η converge to a common limit.

If, in particular, we set $c_n = p_n/b_n$ where p_n is real and positive we have the following result.

THEOREM 4. *The inequalities* (14) *and* (16) *of Theorem* 3 *may be replaced by the following inequalities:*

$$(14') \qquad p_1 \geqq 1, \left| \frac{a_n}{b_{n-1}b_n} \right| \leqq \frac{p_n - 1}{p_n p_{n-1}}, \quad (n = 2, 3, 4, \cdots);$$

$$(16') \qquad \left| \frac{\alpha_{n-1}b_n}{\beta_n x_{n+1}} \right| \geqq s p_n,$$

respectively, where $p_1, p_2, p_3, \cdots$ *are real numbers.*

The special case $p_n = 2$ is important. In this case we readily obtain the following theorem.[7]

THEOREM 5. *If the* x_n, y_n *in* $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$, $(x_n \neq 0)$, *are functions of any variables, then* ξ *converges uniformly over the region determined by the inequalities*

$$\left|\frac{a_1}{b_1}\right| \leqq C, \quad \left|\frac{a_n}{b_{n-1}b_n}\right| \leqq 1/4, \quad (n = 2, 3, 4, \cdots),$$
$$\left|\frac{\alpha_{n-1}b_n}{\beta_n x_{n+1}}\right| \geqq 2s, \quad n > N,$$

where C, N, *and* s *are constants and* $s > 1$.

4. *The transformation* T_1. Let

$$\alpha_{2n} = 1 - g_n y_{2n+1}, \quad \alpha_{2n+1} = -y_{2n+2}/g_n, \quad \beta_{2n} = g_n, \quad \beta_{2n+1} = 1/g_n, \qquad (n = 0, 1, 2, \cdots),$$

where

$$g_n = g_0 - (y_2 + y_4 + \cdots + y_{2n}); \qquad (n = 1, 2, 3, \cdots),$$

and g_0 is any number $\neq 0$ so chosen that $g_n \neq 0$, $n = 1, 2, 3, \cdots$. If $x_n = 1$, the transformation T_1 is as follows:

$$(T_1) \qquad \begin{aligned} &b_0 = y_0 + g_0, \quad a_1 = -1, \quad a_n = 1, \quad n > 1; \\ &b_1 = 1/g_0, \quad b_{2n} = y_{2n-1}g^2_{n-1}, \\ &b_{2n+1} = y_{2n}/g_n g_{n-1}, \end{aligned} \qquad (n = 1, 2, 3, \cdots).$$

Here

$$(17) \qquad A_{2n+1}/B_{2n+1} = X_{2n}/Y_{2n}, \qquad (n = 0, 1, 2, \cdots),$$

so that Theorem 5 in this case takes the following form.

THEOREM 6. *Let* $y_0, y_1, y_2, \cdots$ *be functions of any variables, and let* $g_0 \neq 0$ *be so chosen that* $g_n = g_0 - (y_2 + y_4 + \cdots + y_{2n}) \neq 0$, $n = 1, 2, 3, \cdots$. *Set*

$$t_1 = -g_0, \quad t_2 = \frac{1}{y_1 g_0}, \quad t_{2n} = \frac{g_{n-2}}{y_{2n-1}y_{2n-2}g_{n-1}}, \quad n > 1, \quad t_{2n+1} = \frac{g_n}{y_{2n}y_{2n-1}g_{n-1}}, \quad n > 0;$$
$$k_n = \frac{1}{g_n y_{2n+1}}, \quad n > 0.$$

[7] Perron, *ibid.*, pp. 260-262.

Then the continued fraction $(1, 1, 1, \cdots; y_0, y_1, y_2, \cdots)$ *converges uniformly over the region determined by the inequalities:*

$$| t_1 | < C, \quad | t_n | \leqq 1/4, \qquad (n = 2, 3, 4, \cdots)$$
$$| t_{2n+2}/(k_n - 1) | \leqq 1/2s, \quad n > N,$$

where C, N, *and* s *are constants and* $s > 1$.

As an application we have this theorem:

THEOREM 7. *Let* $h_0, h_1, h_2, \cdots$ *be real or complex numbers* $h_i \neq 0$, *and let* $q_n = q_0 - (h_2 + h_4 + \cdots + h_{2n})$, *where* $q_0 \neq 0$ *is so chosen that* $q_n \neq 0$, $n = 1, 2, 3, \cdots$. *Put*

$$w_{2n} = \frac{q_{n-2}}{q_{n-1}h_{2n-1}h_{2n-2}}, \ n > 1, \qquad w_{2n+1} = \frac{q_n}{q_{n-1}h_{2n}h_{2n-1}}, \ n > 0,$$

and suppose $\lim w_n = 0$. *Then the continued fraction*

$$H = h_0 + z/h_1 + z/h_2 + \cdots$$

has the following properties.

(a) *The sequence of even convergents represents a meromorphic function* $F(z)$ *and converges uniformly over every closed bounded region* R *containing no poles of* $F(z)$.

(b) *Let* $r_n = q_n h_{2n+1}$, *and let* R' *be an arbitrary closed bounded region containing none of the poles of* $F(z)$, *and no limit points of the sequence* $\{r_n\}$. *Then the sequence of odd convergents converges uniformly over* R' *to* $F(z)$.

This theorem is a notable extension of an earlier theroem [8] in which was required the absolute convergence of the series Σw_n.

Write H in the form

$$H' = (1, 1, 1, \cdots; h_0, h_1 z', h_2, h_3 z', \cdots), \qquad z' = 1/z.$$

Then $T_1 H'$ can be put in the form

$$\eta = (w_1 z, w_2 z, \cdots; h_0 + q_0, 1, 1, \cdots).$$

Since $\lim w_n = 0$, η represents a meromorphic function $F(z)$ and converges uniformly over R. By (12) the same is true of the sequence of even convergents of H. This proves (a).

[8] Wall, *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 946-952.

To prove (b), apply Theorem 6 to the continued fraction

$$\xi_{\mu,\nu}=h_{2\mu}+\frac{1}{\nu z' h_{2\mu+1}}+\frac{1}{h_{2\mu+2}}+\frac{1}{z' h_{2\mu+3}}+\cdots,\ z'=1/z,$$

taking $g_0=q_0-(h_2+h_4+\cdots+h_{2\mu})$. Then

$$t_2=\frac{z}{\nu q_\mu h_{2\mu+1}},\quad t_\nu=zw_{2\mu+\nu},\qquad (\nu=3,4,5,\cdots).$$

If we now assign to μ and then to ν sufficiently large values, we will have

$$|t_\nu|\leqq 1/4,\qquad (\nu=2,3,4,\cdots),$$

over a circle K with center $z=0$ and a sufficiently large radius to include the region R'. Also, over R':

$$\left|\frac{t_{2n+2}}{k_n-1}\right|=\left|\frac{zw_{2\mu+2n+2}}{\dfrac{z}{r_{n+\mu}}-1}\right|\leqq\frac{1}{2s},$$

if $n>$ some N, and s is a constant >1. Hence, by Theorem 6, $\xi_{\mu,\nu}$ converges uniformly over R' to a function $f(z)$ which is analytic over K, and is non-rational. Consequently,

$$\frac{1}{h_{2\mu+2}}+\frac{1}{h_{2\mu+3}z'}+\frac{1}{h_{2\mu+4}}+\cdots$$

converges over R' to the value

$$f_1(z)=\frac{1}{f(z)-h_{2\mu}}-\nu z' h_{2\mu+1},$$

and H converges to

$$F(z)=\frac{X_{2\mu+1}+f_1(z)X_{2\mu}}{Y_{2\mu+1}+f_1(z)Y_{2\mu}},$$

which is analytic, except possibly for poles, inasmuch as $f_1(z)$ is non-rational. The convergence is clearly uniform over R'.

THEOREM 8. *Let* $(1,1,1,\cdots;\ 0,y_1,y_2,\cdots)=\xi$ *be a continued fraction with arbitrary partial denominators* $y_1,y_2,y_3,\cdots$; *and suppose the series*

$$(18)\qquad \sum_{n=1}^{\infty} y_{2n+1}(y_2+y_4+\cdots+y_{2n})^2,$$

$$\sum_{n=m}^{\infty}\left[\frac{1}{y_2+y_4+\cdots+y_{2n-2}}-\frac{1}{y_2+y_4+\cdots+y_{2n}}\right],\ m>1,$$

converge absolutely, and that

$$(19)\qquad \lim_{n=\infty}|y_2+y_4+\cdots+y_{2n}|=\infty.$$

Then there exist two numbers F, G, where $|F| + |G| \neq 0$, such that

$$(20)\qquad \begin{aligned} \lim_{n=\infty} \frac{X_{2n}}{y_2 + y_4 + \cdots + y_{2n}} &= \lim_{n=\infty} X_{2n+1} = F, \\ \lim_{n=\infty} \frac{Y_{2n}}{y_2 + y_4 + \cdots + y_{2n}} &= \lim_{n=\infty} Y_{2n+1} = G. \end{aligned}$$

If $y_1, y_2, y_3, \cdots$ are analytic functions of z over a region R in which (18) *converge absolutely and uniformly, then F, G are analytic functions over R,* and *ξ converges to F/G, wherever $G \neq 0$.*

If (18) *converge absolutely and* (19) *fails to hold, then ξ diverges.*

To prove the theorem, apply the transformation T_1 to ξ, and obtain

$$(21)\qquad \frac{X_{2n}}{g_n} = A_{2n+1}, \qquad X_{2n+1} = \frac{A_{2n}}{g_n} - (1 - g_n y_{2n+1}) A_{2n+1},$$

with like expressions for the Y_n. Now from the absolute convergence of (18) it follows that the series

$$(22)\qquad \sum \frac{y_{2n}}{g_{n-1} g_n}, \qquad \sum y_{2n+1} g_n^2$$

converge absolutely, and hence by a theorem of von Koch:[9]

$$(23)\qquad \begin{aligned} \lim A_{2n+1} &= F, \qquad & \lim A_{2n} &= F_0, \\ \lim B_{2n+1} &= G, \qquad & \lim B_{2n} &= G_0, \end{aligned}$$

where $FG_0 - F_0 G = 1$, so that $|F| + |G| \neq 0$. If (18), and hence (22), converge absolutely and uniformly over R, then F, G, F_0, G_0 are analytic functions over R. The equations (20) now follow at once from (21), (23), and the hypothesis.

It is obvious that when (18) converge absolutely and (19) fails to hold, then $\Sigma |y_n|$ converges, and hence ξ diverges.

Remark. The second series (18) evidently converges when (19) holds. If we could remove the requirement of absolute convergence of this series, this theorem would include the following theorem of Hamburger:[10] *Let* $(1, 1, 1, \cdots; 0, a_1 z, a_2, a_3 z, \cdots)$ *be a continued fraction in which the a_n are real and $\neq 0$, and $a_{2n+1} > 0$. Then, if the series $\Sigma a_{2n+1}(a_2 + a_4 + \cdots + a_{2n})^2$ converges and* $\lim |a_2 + a_4 + \cdots + a_{2n}| = \infty$, *the continued fraction converges to a function which is analytic except for poles.*

---

[9] Von Koch, *Bulletin de la Société Mathématique de France*, vol. 23 (1895).

[10] H. Hamburger, *Mathematische Annalen*, vol. 82, pp. 120-164, 168-187.

5. *Transformations which interchange two convergents.* We proceed to obtain a transformation $T = W_m$ which interchanges the m-th and $(m+1)$-th convergents of $\xi = (x_1, x_2, \cdots; y_0, y_1, \cdots)$. The interchange of any two convergents can be effected by repeated application of W_m with properly chosen values of m. We find that the formulas for the transformation W_m are as follows:

$$(W_0)\quad \begin{cases} b_0 = y_0 + \dfrac{x_1}{y_1},\ a_1 = -\dfrac{x_1}{y_1{}^2},\ a_2 = y_2,\ a_3 = -\dfrac{x_2 x_3}{y_2}, \\ b_1 = \dfrac{1}{y_1},\ b_2 = x_2,\ b_3 = \dfrac{y_2 y_3 + x_3}{y_2};\ a_n = x_n,\ b_n = y_n, \\ \qquad\qquad\qquad\qquad\qquad\qquad (n = 4, 5, 6, \cdots), \end{cases}$$

$$(W_m)\quad \begin{cases} a_m = x_m y_{m+1},\ a_{m+1} = -x_{m+1}/y_{m+1},\ a_{m+2} = y_{m+2}, \\ a_{m+3} = -x_{m+2}x_{m+3}/y_{m+2},\ b_m = y_m y_{m+1} + x_{m+1}, \\ b_{m+1} = 1/y_{m+1},\ b_{m+2} = x_{m+2},\ b_{m+3} = \dfrac{y_{m+2}y_{m+3} + x_{m+3}}{y_{m+2}}, \\ a_n = x_n, \qquad\qquad (n \neq m,\ m+1,\ m+2,\ m+3) \\ b_n = y_n \qquad\qquad\qquad (m = 1, 2, 3, \cdots). \end{cases}$$

This transformation is valid if $y_{m+1}y_{m+2} \neq 0$. If $b_n, y_n \neq 0$, $n = 1, 2, 3, \cdots$, and we set

$$\begin{aligned} t_1 &= x_1/y_1, \quad t_n = x_n/y_{n-1}y_n, \\ t'_1 &= a_1/b_1, \quad t'_n = a_n/b_{n-1}b_n, \end{aligned} \qquad (n = 2, 3, 4, \cdots),$$

the transformation becomes:

$$(W_0)\quad \begin{cases} b_0 = y_0 + t_1,\ t'_1 = -t_1,\ t'_2 = 1/t_2,\ t'_3 = \dfrac{-t_3}{1+t_3},\ t'_4 = \dfrac{t_4}{1+t_3}, \\ t'_n = t_n, \qquad\qquad (n = 5, 6, 7, \cdots); \end{cases}$$

$$(W_m)\quad \begin{cases} t'_m = \dfrac{t_m}{1+t_{m+1}},\ t'_{m+1} = -\dfrac{t_{m+1}}{1+t_{m+1}},\ t'_{m+2} = \dfrac{1}{t_{m+2}}, \\ \quad t'_{m+3} = \dfrac{-t_{m+3}}{1+t_{m+3}},\ t'_{m+4} = \dfrac{t_{m+4}}{1+t_{m+3}}, \\ \quad t'_n = t_n \quad (n = 1, 2, 3, \cdots, m-1,\ m+5,\ m+6, \cdots), \\ \qquad\qquad\qquad (m = 1, 2, 3, \cdots). \end{cases}$$

Let T_2 denote the transformation which results by applying in succession $W_0, W_2, W_4, \cdots$. Then the formulas for T_2 are:

$$(T_2)\quad \begin{cases} b_0 = y_0 + t_1,\ t'_1 = -t_1,\ t'_2 = \dfrac{1+t_3}{t_2}, \\ \qquad t'_{2n} = \dfrac{(1+t_{2n-1})(1+t_{2n+1})}{t_{2n}}, \qquad (n = 2, 3, 4, \cdots), \\ \qquad t'_{2n+1} = t_{2n+1}, \qquad\qquad (n = 1, 2, 3, \cdots). \end{cases}$$

If we apply in succession $W_1, W_4, W_7, W_{10}, \cdots$, we obtain:

$$(T_3) \quad \begin{cases} b_0 = y_0,\ t'_1 = \dfrac{t_1}{1+t_2},\ t'_2 = \dfrac{-t_2}{1+t_2}, \\ t'_{3n} = \dfrac{1}{t_{3n}},\ t'_{3n+1} = \dfrac{-t_{3n+1}}{1+t_{3n+1}+t_{3n+2}},\ t'_{3n+2} = \dfrac{-t_{3n+2}}{1+t_{3n+1}+t_{3n+2}}, \\ \qquad\qquad (n=1,2,\cdots). \end{cases}$$

These transformations furnish immediate extension of Pringsheim's criteria. For example, from W_0 we obtain this result:

THEOREM 9. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, 1, \cdots)$ $(t_n \neq 0)$ *converges if*

$$|t_2| \geqq 4, \quad |t_3| \leqq 1/5, \quad |t_4| \leqq 1/5, \quad |t_n| \leqq 1/4, \quad (n=5,6,7,\cdots).$$

From T_2 we obtain the theorem:

THEOREM 10. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, 1, \cdots)$ $(t_n \neq 0)$ *converges if there exist real numbers* $p_1, p_2, p_3, \cdots, p_i > 1$, *such that*

$$|t_{2n+1}| \leqq \frac{p_{2n+1}-1}{p_{2n}p_{2n+1}}, \qquad |t_{2n}| \geqq M_{2n}, \qquad (n=1,2,3,\cdots)$$

where

$$M_2 = \frac{p_1(p_2p_3+p_3-1)}{p_3(p_2-1)}, \quad M_{2n} = \frac{(p_{2n-2}p_{2n-1}+p_{2n-1}-1)(p_{2n}p_{2n+1}+p_{2n+1}-1)}{(p_{2n}-1)(p_{2n-2}p_{2n+1})} \quad (n>1).$$

If $p_n = 2$ this gives the first theorem mentioned at the end of the introduction.

From T_3 we have:

THEOREM 11. *The continued fraction* $(t_1, t_2, t_3, \cdots; y_0, 1, \cdots)$ $(t_n \neq 0)$ *converges if*

$$|t_2| \leqq 1/5, \quad |t_{3n}| \geqq 4, \quad |t_{3n+1}| \leqq 1/6, \quad |t_{3n+2}| \leqq 1/6, \quad (n=1,2,3,\cdots).$$

We shall now apply the transformation W_m to prove the following theorem mentioned at the end of § 1.

THEOREM 12. *Let* $q_1, q_2, q_3, \cdots$ *be constants* $\neq 0$, *and let*

$$\lim_{v=\infty} q_{n'_v} = \infty, \qquad \lim_{v=\infty} q_{n''_v} = 0,$$

where $B = (n_1' < n_2' < \cdots)$ *and* $A = (n_1'' < n_2'' < \cdots)$ *are mutually exclusive sets whose sum is the set of natural numbers* $1, 2, 3, \cdots$; *and suppose that if* n *is in* B, *then* $n - 1$ *and* $n + 1$ *are in* A. *Then, corresponding to every bounded region* R *of the complex* z-*plane whose distance from the origin is* > 0, *there exists an index* N *such that if* $n \geqq N$ *the continued fraction*

$$1 + \frac{q_n z}{1} + \frac{q_{n+1} z}{1} + \cdots \tag{24}$$

converges uniformly over R.

Proof. Choose N sufficiently large to insure that, over R,

$$|q_\nu z| \leqq \delta,\ \nu \geqq N,\ \nu \text{ in } A;\ \left|\frac{1}{q_\nu z}\right| \leqq \delta,\ \nu \geqq N,\ \nu \text{ in } B,$$

where δ is a properly chosen small positive number. Let $n'_\mu, n'_{\mu+1} \cdots$ be the numbers in B which are $> n \geqq N$, and apply to (24) in succession the transformations

$$W_{n'_i - 2}, \qquad (i = \mu, \mu + 1, \mu + 2, \cdots).$$

Then in the resulting continued fraction, $(t'_1, t'_2, t'_3, \cdots;\ 1, 1, 1, \cdots)$, we shall have, over R,

$$|t'_n| \leqq 1/4, \qquad (n = 2, 3, 4, \cdots)$$

and hence, (24) converges uniformly over R.

6. *The transformation* T_4. This transformation replaces the $2n$-th convergent by the $(2n + 1)$-th, $(n = 0, 1, 2, \cdots)$. The formulas are found to be

$$(T_4) \qquad \begin{aligned} &b_0 = y_0 + \frac{t_1}{1 + t_2}, \qquad t'_1 = \frac{t_1 t_2}{1 + t_2}, \\ &t'_{2n} = \frac{(1 + t_{2n})(1 + t_{2n+2})}{t_{2n+1}}, \qquad t'_{2n+1} = t_{2n+2}, \qquad (n = 1, 2, 3, \cdots). \end{aligned}$$

If we follow the transformation T_2 by T_4 we shall obtain

$$b_0 = y_0 - \frac{t_1 t_2}{1 + t_2 + t_3}, \quad t'_1 = \frac{-t_1(1 + t_3)}{1 + t_2 + t_3},$$

$$T'_2 = \frac{\left(1 + \frac{1 + t_3}{t_2}\right)\left(1 + \frac{(1 + t_3)(1 + t_5)}{t_4}\right)}{t_3},$$

$$(T_4T_2) \quad t'_{2n} = \frac{\left(1 + \dfrac{(1+t_{2n-1})(1+t_{2n+1})}{t_{2n}}\right)\left(1 + \dfrac{(1+t_{2n+1})(1+t_{2n+3})}{t_{2n+2}}\right)}{t_{2n+1}}, \qquad (n = 2, 3, 4, \cdots),$$

$$t'_{2n+1} = \frac{(1+t_{2n+1})(1+t_{2n+3})}{t_{2n+2}}, \qquad (n = 1, 2, 3, \cdots).$$

From this, Pringsheim's criterion gives the following theorem.

THEOREM 13. *Let $p_1, p_2, p_3, \cdots$ be real and > 1, and let $M_n = \dfrac{p_n - 1}{p_n p_{n-1}}$, $n > 1$, $M_1 > 0$,*

$$\bar{M}_{2n} = \frac{(1+M_{2n-1})(1+M_{2n+1})}{M_{2n}}, \qquad (n = 1, 2, 3, \cdots).$$

Let $M'_{2n} \geqq \bar{M}_{2n}$, and set $\bar{M}'_{2n} = (1+M'_{2n})(1+M'_{2n-2})/M_{2n-1}$, $n = 1, 2, 3, \cdots$ $M'_0 = 0$. Then the continued fraction $(x_1, x_2, \cdots ; y_0, 1, 1, \cdots)$ converges if

$$(25) \qquad M'_{2n} \geqq |t_{2n+1}| \geqq \bar{M}_{2n}, \qquad |t_{2n}| \geqq \bar{M}'_{2n}, \qquad (n = 1, 2, 3, \cdots).$$

If $p_n = 2$, $M'_{2n} = M$, then (25) becomes

$$(25') \qquad M \geqq |t_{2n+1}| \geqq 25/4, \quad |t_{2n}| \geqq 4(1+M)^2, \qquad (n = 1, 2, 3, \cdots).$$

If $p_n = 2$ and $\lim M'_{2n} = \infty$, then $\lim |t_{2n}/t_{2n+1}| = \infty$.

By application of the transformations $T_2, T_4T_2, T_4{}^2T_2, \cdots$ we obtain

THEOREM 14. *Let $h_1, h_2, h_3, \cdots$ be any numbers $\neq 0$ such that for some $k \geqq 0$:*

$$(26) \qquad \lim_{n=\infty} h_{2n}h_{2n+2k-1} = 0, \qquad \lim_{n=\infty} h_{2n}h_{2n+2k+1} = \infty.$$

Then, corresponding to every bounded region R of the complex z-plane whose distance from the origin is positive, there exists an index N such that if $n \geqq N$ the continued fraction

$$(27) \qquad h_{2n} + \frac{z}{h_{2n+1}} + \frac{z}{h_{2n+2}} + \cdots$$

converges uniformly over R.

To prove the theorem, let $q_1 = 1/h_1$, $q_n = 1/h_{n-1}h_n$, $n > 1$. Then (27) takes the form

$$h_{2n}\left[1 + \frac{q_{2n+1}z}{1} + \frac{q_{2n+2}z}{1} + \cdots\right] = h_{2n}\xi_n;$$

and $T_2\xi_n = (t'_1, t'_2, \cdots; b_0, 1, 1, \cdots)$, where

$$t'_2 = \frac{1 + q_{2n+1}z}{q_{2n}z}, \qquad t'_{2n} = \frac{(1 + q_{2n-1}z)(1 + q_{2n+1}z)}{q_{2n}z},$$
$$n > 1, \; t'_{2n+1} = q_{2n+1}z, \; n > 0.$$

Hence, if $\lim q_{2n+1} = \lim (1/h_{2n}h_{2n+1}) = 0$, $\lim q_{2n} = \lim (1/h_{2n}h_{2n+1}) = \infty$, it is clear that there exists an index N such that if $n \geqq N$, $|t'_\nu| \leqq 1/4$, $\nu = 2, 3, 4, \cdots$, over R, so that ξ_n, and therefore (27), converge uniformly over R. This proves the theorem for the case $k = 0$.

In like manner, when $k = 1, 2, \cdots$ the theorem may be proved by using the transformations $T_4T_2, T_4^2T_2, \cdots$, respectively. The proof for all k can be readily accomplished by mathematical induction.

Example. Let $(z, z, z, \cdots; 0, h_1, h_2, h_3, \cdots)$ be a Stieltjes continued fraction in which the h_n are real and positive. Then if (26) holds, Σh_ν diverges and hence the continued fraction converges, except along the whole or a part of the negative half of the real axis. Inasmuch as (27) is an analytic function of z over R, if n is sufficiently large, *it follows that our Stieltjes continued fraction represents a function whose only singularities in the region R are poles.* The same argument can be used if the h_n are real and $\neq 0$ and $h_{2n+1} > 0$. The conditions of Theorem 14 are met if

$$h_{2n} = r^{n^2}u_{2n}, \qquad h_{2n-1} = \frac{u_{2n-1}}{r^{(n-k-1/2)^2}}, \qquad (n = 1, 2, 3, \cdots),$$

where $0 < r < 1$, and $u_1, u_2, u_3, \cdots$ are any constants which are bounded and bounded away from 0.

HARVARD UNIVERSITY
AND
NORTHWESTERN UNIVERSITY.

THE FORM $2wx + xy + yz + zu + ux$.

By E. T. Bell.

1. In a recent paper[1] I gave a set of 25 arithmetical identities equivalent to all those of a certain sort between doubly periodic functions of the first and second kinds, all referring to the arithmetical form $xy + zw$. From the same identities another set concerning the form $wx + xy + yz$, or the like with numerical coefficients > 1, can be derived, and these in turn yield interesting results concerning numbers of representations in quinary quadratic forms. To illustrate one method of obtaining the new identities we shall prove the following.

Theorem 1. *Let t, τ, w, x, y, z be integers > 0 such that, for m odd and constant,*

$$(1) \qquad m = t\tau = wx + xy + 2yz,$$

and let $f(u)$ be any even function of u which is defined for integer values of u,

$$(2) \qquad f(u) = f(-u),\ u \text{ integral.}$$

Then

$$(3) \quad \Sigma\,[(x+z)f(w-y) - zf(w+y)] = \Sigma \sum_{r=1}^{(\tau-1)/2} [rf(t) - tf(\tau + 2 - 4r)],$$

the Σ on the left referring to all (w, x, y, z), that on the right to all (t, τ) satisfying (1).

Let $N[n = F; *]$ denote the number of representations of n in the form F, the integer variables in F being restricted by the conditions *. We shall prove from (3),

Theorem 2.

$$\zeta_2(m) - (4m+1)\zeta_0(m) + 4\zeta_1(m)$$
$$= 8N[m = 2wx + xy + yz + zu + ux;\ w, x, y, z > 0;\ u \geqq 0],$$

in which m is odd, and $\zeta_r(m)$ denotes the sum of the r-th powers of all the divisors of m (so that $\zeta_0(m)$ is the number of divisors).

This result is unusual in that it expresses the number of representations in an appropriately restricted indefinite form in five variables in terms of the

[1] *American Journal of Mathematics*, vol. 57 (1935), pp. 245-253.

divisors of the integer represented. There seems to be but one other similar result in the literature, stated in 1866 by Liouville (a proof of which will be published elsewhere).

2. In the formula (II), p. 249 of the paper cited we may choose $|y| f(x)$ for the function of x, y there, where $f(x)$ is as in (2). We get

$$(4)\quad \sum [(\tau_1 + \tau_2) f(t_1 - 2t_2) - |\tau_1 - \tau_2| f(t_1 + 2t_2)] = 2 \sum \sum_{r=1}^{(\tau-1)/2} r f(t),$$

the sums referring to all $(t_1, \tau_1, t_2, \tau_2)$, (t, τ) from

$$(5)\quad m = t\tau = t_1\tau_1 + 2t_2\tau_2,$$

in which all letters denote integers > 0, m, τ_2 are odd (and hence also t_1, τ_1).

Considering the second term in (4) we distinguish the cases $\tau_1 > \tau_2$, $\tau_1 < \tau_2$, $\tau_1 = \tau_2$, and make the corresponding changes in (5):

$$(6)\quad \begin{aligned} &\tau_1 > \tau_2;\ \tau_1 = \tau_2 + 2x,\ x > 0;\\ &m = t_1(\tau_2 + 2x) + 2t_2\tau_2,\ t_1, \tau_2 \text{ odd};\end{aligned}$$

$$(7)\quad \begin{aligned} &\tau_2 > \tau_1;\ \tau_2 = \tau_1 + 2y,\ y > 0;\\ &m = t_1\tau_1 + 2t_2(\tau_1 + 2y),\ t_1, \tau_1 \text{ odd};\end{aligned}$$

$$(8)\quad \tau_1 = \tau_2;\ m = t\tau = \tau_1(t_1 + 2t_2).$$

Solving (8), we have $\tau_1 = \tau$, $t_1 + 2t_2 = t$. For (t, τ) a fixed pair of divisors of m there are the solutions

$$\{t_1, t_2\} = \{2j - 1, (t - 2j + 1)/2\}, \qquad (j = 1, \cdots, (t-1)/2).$$

Hence (8) contributes

$$(9)\quad 2 \sum \tau \sum_{j=1}^{(\tau-1)/2} f(t - 4j + 2)$$

to the left of (4), Σ referring to all (t, τ) in (5). We may interchange t, τ in (9). Thus (4) becomes

$$(10)\quad \begin{aligned} &\sum_6 [(\tau_2 + x) f(t_1 - 2t_2) - x f(t_1 + 2t_2)]\\ &+ \sum_7 [(\tau_1 + y) f(t_1 - 2t_2) - y f(t_1 + 2t_2)]\\ &= \sum \sum_{r=1}^{(\tau-1)/2} [r f(t) - t f(\tau + 2 - 4r)],\end{aligned}$$

where $\sum_6$, $\sum_7$ refer to the forms in (6), (7). Both forms are included in

$$m = w(\tau_2 + 2x) + v\tau_2,\ \tau_2 \text{ odd},\ w \not\equiv v \bmod 2.$$

Since m is odd, the conditions on τ_2, w, v are automatically satisfied and may be suppressed. Using the condition (2) we have

$$\sum_6 + \sum_7 = \sum [(y+x)f(w-v) - xf(w+v)],$$

the sum referring to all sets (v, y, w, x) of integers > 0 such that

$$m = vy + yw + 2wx.$$

The change of notation $(v, y, w, x) \rightarrow (w, x, y, z)$ then gives the result in Theorem 1 from (10).

3. To prove Theorem 2 take $f(x) = 1$ for all integer values of x in (3). A simple reduction gives

$$8 \sum x = \zeta_2(m) - (4m+1)\zeta_0(m) + 4\zeta_1(m),$$

the sum extending over all solutions of (1). But

$$x = N[x = x_1 + x_2;\ x_1 > 0,\ x_2 \geqq 0].$$

Hence we get the result stated by the change of notation

$$(z, y, x_1, w, x_2) \rightarrow (w, x, y, z, u).$$

CALIFORNIA INSTITUTE OF TECHNOLOGY.

REDUCIBLE EXCEPTIONAL CURVES.

By Patrick Du Val.

The following research was undertaken and to a great extent carried out independently of that of Barber and Zariski on the same topic.[1] It may be regarded as a continuation, extension, and I hope a simplification, of the results of those authors. They only consider in much detail reducible curves transformable into the neighbourhoods of a sequence of points along a single branch; but I shall shew that very simple and elegant results hold for the neighbourhoods of a set of points grouped round each other in any way.

The essential idea for which I may refer the reader to the paper already mentioned is that each point of the set has two neighbourhoods, a total neighbourhood which is a curve (generally reducible) of virtual genus 0 and virtual grade -1, and a diminished neighbourhood which is an irreducible rational curve, of virtual grade generally < -1. The total neighbourhood of a point not in the neighbourhood of any other, i. e. of what we may call an original point, is what is subtracted from any linear system not having a base point at the point in question, by giving it a simple base point there; the diminished neighbourhood is what is left of this after those of its points which belong to the set in question have been transformed into curves. It is evident from this that the diminished neighbourhood of the point is equal to the total neighbourhood, less the total neighbourhoods of certain other points of the set. For a point which is not original we apply the definition by first transforming all its ancestor points into curves, when it becomes original.

It is evident that the points can be arranged in such an order that none is in a neighbourhood of any which follows it. This we shall call a standard order. It is unique if and only if the points are consecutive on a single branch.

We shall suppose the set of points to be n in number, and call them $O_1, \cdots, O_n$, in a standard order. The total and diminished neighbourhoods of O_a we shall call Q_a, L_a, respectively; the set of curves Q_a, arranged as a matrix of one row and n columns, will be denoted by $\mathbf{Q}$, the same set arranged as a matrix of one column and n rows by $\tilde{\mathbf{Q}}$; $\mathbf{L}$ and $\tilde{\mathbf{L}}$ will have similar meanings. (Throughout, clarendon type will indicate matrices, and the symbol $\tilde{}$ the transposition of a matrix, or interchange of rows and columns.) From what has been said it is clear that there is a relation

[1] S. F. Barber and Oscar Zariski, "Reducible exceptional curves of the first kind," *American Journal of Mathematics*, vol. 57 (1935), pp. 119-141.

(1) $$\bar{L} = m\bar{Q},$$

where m is a numerical square matrix, in which (i) every element of the main diagonal is 1, (ii) every element to the left of the diagonal is 0, and (iii) certain elements to the right of the diagonal are -1, and the rest 0. The determinant $|\, m \,| = 1$, so that we can also write

(2) $$\bar{Q} = m^{-1}\bar{L}.$$

We shall indicate the virtual intersection number of two curves or linear systems as a dot product; since all the curves Q_α have virtual grade -1, and no virtual intersections with each other, they have the intersection matrix

$$\bar{Q} \cdot Q = -E,$$

where E indicates the unit or identical matrix; hence by (1) that of the curves L_α is

(3) $$\bar{L} \cdot L = m\bar{Q} \cdot Q\bar{m} = -m\bar{m} = n$$

say; the elements of the numerical matrices m, n, we shall write $m_{\alpha\beta}$, $n_{\alpha\beta}$.

If $m_{\alpha\beta} = -1$, the point O_β will be called proximate to O_α. (It is easy to see that this is equivalent to the definition of proximate points taken by Barber and Zariski from Enriques.) Since the curves L_α are irreducible, $n_{\alpha\beta}$ cannot be negative for $\alpha \neq \beta$, and we deduce at once the following results from (3):

If any point is proximate to two others, then one of these is proximate to the other; for if we had $m_{\alpha\gamma} = m_{\beta\gamma} = -1$, $m_{\alpha\beta} = 0$, $\alpha < \beta < \gamma$, then $n_{\alpha\beta}$ would be negative; similarly

No two points can be proximate to the same two others, since if we had $m_{\alpha\gamma} = m_{\beta\gamma} = m_{\alpha\delta} = m_{\beta\delta} = -1$, $\alpha < \beta < \gamma < \delta$, then even with $m_{\alpha\beta} = -1$, $n_{\alpha\beta}$ would be negative; finally

No point can be proximate to more than two others; for if we had $m_{\alpha\delta} = m_{\beta\delta} = m_{\gamma\delta} = -1$, $\alpha < \beta < \gamma < \delta$, then also $m_{\alpha\gamma} = m_{\beta\gamma} = m_{\alpha\beta} = -1$, and $n_{\alpha\beta}$ would be negative.

If $|\, K \,|$ be the canonical system on the surface, let $L \cdot K = s$, $Q \cdot K = t$; since the curves L_α, Q_α are all of genus 0, $t_\alpha = -1$, and $s_\alpha = -(2 + n_{\alpha\alpha})$. We must of course have

(4) $$\bar{s} = m\bar{t};$$

but this also follows from (3), since from the form of m we have

$$-\sum_\beta (m_{\alpha\beta}) = \sum_\beta (m^2{}_{\alpha\beta}) - 2.$$

We now proceed to prove the general converse of these properties, or the generalisation of the well-known theorem:

(5) *If an irreducible curve has genus* 0 *and grade* -1, *then there exist regular systems, free from unassigned base points, on whose projective models the curve appears as the neighbourhood of a simple point.*

The theorem we shall prove is as follows:

(6) *If a set of irreducible curves all have genus* 0, *and if (for a suitable ordering of the curves) their intersection matrix can be put into the form* $-\boldsymbol{m}\bar{\boldsymbol{m}}$, *where* $\boldsymbol{m}$ *has the properties* (i), (ii), (iii), *above, then there exist regular systems, free from unassigned base points, on whose projective models the curves appear as the diminished neighbourhoods of a set of simple points.*

This we shall prove by induction, assuming it to be true for a set of $n-1$ curves, and hence proving it to be true for a set of n. The Theorem (5) itself provides the basis of induction, being precisely the case $n=1$.

We note first, since the last row of $\boldsymbol{m}$ consists entirely of 0's, except for a 1 in the last place, that the last column of $\boldsymbol{n}$ is the same as that of $\boldsymbol{m}$ with the sign changed; and in particular $n_{nn}=-1$; thus L_n is identical with Q_n, and is an exceptional curve as it stands. We can therefore, by (5), transform the surface Φ on which the curves are into a surface Φ^* on which L_n appears as the neighbourhood of a simple point O^*_n; the remaining curves L_a appear as a set of $n-1$ curves L^*_a, where L^*_a is the image of L_a or of L_a+L_n according as $n_{an}=0$ or 1. The intersection matrix of the curves $\boldsymbol{L}^*$ is $\boldsymbol{n}^*$, obtained from $\boldsymbol{n}$ by omitting the last row and column, and increasing by 1 each element $n_{a\beta}$ such that $n_{an}=n_{n\beta}=1$ (this just takes care of the increase of grade of those curves to which L_n is added, and of the possible intersection of two of the new curves in O^*_n). But on account of the identity (save for sign) of the last row and column of $\boldsymbol{n}$ with the last column of $\boldsymbol{m}$, this means that

$$\boldsymbol{n}^*=-\boldsymbol{m}^*\bar{\boldsymbol{m}}^*,$$

where $\boldsymbol{m}^*$ is obtained from $\boldsymbol{m}$ by omitting the last row and column. Thus the curves L^*_a are a set to which the inductive hypothesis is applicable, and there exists a linear system, regular and free from unassigned base points on Φ^*, on whose projective model $\underline{\Phi}$ the curves L^*_a appear as the diminished neighbourhoods of a set of $n-1$ simple points O_a. This system is equally free from unassigned base points on the original surface Φ, and as it certainly

has not a base point at O^*_n the latter appears on $\underline{\Phi}$ as a simple point $\underline{O_n}$; the diminished neighbourhoods of the whole set of n points $\underline{O_a}$ are the images of the original set of curves L_a; thus the theorem is proved.

A square matrix is called reducible, if (possibly after a rearrangement of the rows, and corresponding rearrangement of the columns) it consists of a number of smaller square matrices arranged corner to corner down the main diagonal, and zeros everywhere else; these smaller matrices are called its constituents. If not reducible it is called irreducible. It is evident that $\boldsymbol{m}$, $\boldsymbol{n}$ are reducible or irreducible together.

We have the theorem:

If $\boldsymbol{m}$ and $\boldsymbol{n}$ are irreducible, the curves L_a are a simply connected tree (i. e., any two of them can be linked in an unique way by a chain of curves of the set, all different, and in which consecutive curves and only those intersect.) This also we prove by induction; for if $\boldsymbol{m}$ is irreducible, so is $\boldsymbol{m}^*$ and hence also $\boldsymbol{n}^*$; on the other hand if $\boldsymbol{n}$ is irreducible, L_n meets either one or two of the other curves, i. e., O^*_n is either a general point of one of the curves L^*_a, or the common point of two of them; and this being so, it is clear that if the curves L^*_a form a simply connected tree, so do the curves L_a.

Another thing that is evident if $\boldsymbol{m}$ and $\boldsymbol{n}$ are irreducible is that O_1 is the only one of the points O_a that is original, since if $n > 1$, O_n is certainly not original, being as we have seen proximate to at least one other point.

On the other hand, it is clear that if $\boldsymbol{m}$ be reducible, and $\boldsymbol{m}^{(i)}$ its constituents, those of $\boldsymbol{n}$ are $\boldsymbol{n}^{(i)} = -\boldsymbol{m}^{(i)}\bar{\boldsymbol{m}}^{(i)}$, and that to each constituent of either corresponds a simply connected tree among the curves L_a, having no intersection with any of the other trees, and a subset of the points O_a consisting of one original point and other points in its various neighbourhoods; the different original points being at distinct points of Φ.

In conclusion, it may be pointed out that a general linear system $|\,C\,|$ on Φ corresponds to a system $|\,C\,|$ on $\underline{\Phi}$ having base points of multiplicity h_a at the points $\underline{O_a}$, where

$$\boldsymbol{h} = C \cdot \boldsymbol{Q} = C \cdot \boldsymbol{L}\bar{\boldsymbol{m}}^{-1},$$

so that the linear system without base points on $\underline{\Phi}$ in which $|\,C\,|$ is totally contained is

$$\begin{aligned} |\,C + C \cdot \boldsymbol{Q}\bar{\boldsymbol{Q}}\,| &= |\,C + C \cdot \boldsymbol{L}\bar{\boldsymbol{m}}^{-1}\boldsymbol{m}^{-1}\bar{\boldsymbol{L}}\,| \\ &= |\,C + C \cdot \boldsymbol{L}\boldsymbol{n}^{-1}\bar{\boldsymbol{L}}\,|. \end{aligned}$$

Note. The criterion for the theorem (6), that the matrix $\boldsymbol{n}$ shall be expressible in the form $-\boldsymbol{m}\bar{\boldsymbol{m}}$ is one that can be directly tested by constructing the matrix $\boldsymbol{m}$ if it exists. For since the last column of $\boldsymbol{m}$ is the

same as that of $\boldsymbol{n}$ with the sign changed, we can observe if $\boldsymbol{n}$ has any column consisting of -1 in the diagonal, and not more than two 1's elsewhere; if it has, we can rearrange the columns (and the rows correspondingly) so that this is the last, and take the last column of $\boldsymbol{m}$ to be the same with change of sign; we then construct the matrix $\boldsymbol{n}^*$ as before, and repeat the process to find the last column but one, and so on; remembering of course that every rearrangement of the rows and columns must be echoed as a rearrangement of the rows of the part of $\boldsymbol{m}$ already built up. If at any stage there is no diagonal element left which is -1, the process breaks down, and $\boldsymbol{m}$ does not exist; if we see a diagonal element -1 with more than two 1's in the same row or column, we can give up, as the process will certainly break down later on.

CAMBRIDGE, ENGLAND.

THEOREMS CONCERNING A BASIS FOR ALL SOLVABLE GROUPS, WITH A LIST OF BASIS GROUPS OF DEGREE < 64.

By A. C. LUNN and J. K. SENIOR.

In previous papers [1] it has been shown that *every solvable group of order* Πx_i $(x_i = p_i^{a_i},\ i = 1 \cdots n)$ *can be expressed and in one way only as a permutation group of degree* Σx_i *with n sets of transitivity of degrees respectively* x_i. Such expressions will here be called ΣU^* expressions. This result directs attention to a special category of permutation groups (to be called U^*-groups) having the following properties: Each U^* is transitive of some degree p^a and of some order Kp^a where K is prime to p, and is a solvable group containing no proper invariant subgroup of order dividing K. Clearly each transitive constituent of a ΣU^* expression is a U^*-group and conversely every U^*-group occurs as a transitive constituent of some ΣU^* expression, e. g. the ΣU^* expression of the abstract group (hereafter called a U-group) with which the U^*-group in question is simply isomorphic. It has been proven that if two U^*-groups are distinct as permutation groups, they cannot be simply isomorphic and hence there is a one to one correspondence between U^*- and U-groups. The category of U-groups is of interest since every solvable abstract group is either a U-group or can be formed from a properly chosen set of U-groups by direct multiplication or by the use of suitable isomorphisms. The U-groups thus constitute a basis, from the members of which all solvable groups may be constructed.

The one to one correspondence between U- and U^*-groups makes it convenient to identify U-groups by their U^*-expressions. A U^*-group of degree p^a and order Kp^a will be called $U^*(p^a, K)$. The purpose of the present paper is to consider certain properties of the U^*-groups and to give a complete list of the $U^*(p^a, K)$ groups where $p^a < 64$, $K > 1$ and $a > 1$. It will appear that where $K = 1$ or $a = 1$ the $U^*(p^a, K)$ groups belong to categories which are already too well known to require relisting.

Certain properties of U^*-groups which are useful in listing all those of given degree are shown in the two following theorems.[2] It has however been

[1] Lunn and Senior, *American Journal of Mathematics*, vol. 56 (1934), pp. 319, 511.

[2] In developing these theorems, we have been greatly assisted by valuable sug-

convenient to state these theorems in a form somewhat more general than that required for the immediate purpose. The following notation will be used.

(1) L and K are two relatively prime integers.
(2) If the order of a solvable group divides L or K, the group will be called respectively an L-group or a K-group.
(3) A solvable group whose order divides LK will be called a V^L-group if it contains no proper invariant K-subgroup; it will be called a V^K-group if it contains no proper invariant L-subgroup.
(4) $A(G)$ is defined as the group of automorphisms of G.
$B(G)$ " " " " " " inner " of G.
$O(G)$ " " " " " " quotient group $A(G)/B(G)$.

Note that

(a) Every L-group is a V^L-group and every K-group is a V^K-group, but not conversely.

(b) A solvable group G of order LK cannot be both a V^L-group and a V^K-group, since at least one of its sets of Sylow subgroups has an intersection invariant under G and of order > 1.

THEOREM I. *If G of order LK is a V^L-group:*

(a) *G uniquely determines each member of a series of characteristic subgroups of G such that $I = G_0 < G_1 \cdots < G_{n-1} < G_n = G$ and such that the quotient group G_{i+1}/G_i is the maximal invariant L-subgroup in G_n/G_i where i is even and the maximal invariant K-subgroup in G_n/G_i where i is odd.* This series will be called the V-series of G, and G (a V^L-group) will be written $V^L{}_n$, if it has $n-1$ proper subgroups in its V-series.

(b) *If G_i is a member of the V-series of G ($= V^L{}_n$) then G_i is a $V^L{}_i$-group and the $i+1$ members of its V-series are identical respectively with the first $i+1$ members of the V-series of G.* Hence the V-series of G may be written

$$I = V^L{}_0 < V^L{}_1 \cdots < V^L{}_{n-1} < V^L{}_n = G.$$

A $V^L{}_{i+a}$ group will be said to be "stepped up" from a certain $V^L{}_i$-group when this $V^L{}_i$-group is the $(i+1)$-th member of the V-series of $V^L{}_{i+a}$.

Proof. All of the invariant L-subgroups of G generate the maximal

---

gestions from Mr. Philip Hall of Cambridge, England, and Mr. Garrett Birkhoff of Cambridge, Mass.

invariant L-subgroup of G. This characteristic subgroup of G is G_1. Since G is solvable, G/G_1 is solvable, and if $G > G_1$, then since G/G_1 contains no invariant L-subgroup, it contains at least one invariant K-subgroup. All the invariant K-subgroups of G/G_1 generate the maximal invariant K-subgroup of G/G_1. To this characteristic subgroup of G/G_1, there corresponds a characteristic subgroup of G which is G_2. By repetition of this argument, the V-series of G is completely determined.

Let F be any invariant subgroup of G. If F is an L-group, then it is also a V^L-group. If F is not an L-group and contains an invariant K-subgroup H, then H (being solvable) must contain at least one prime power subgroup J which is characteristic under H, hence invariant under F, and thus contained in the intersection D of a set of Sylow subgroups of F. This D would be a K-subgroup, characteristic under F and hence invariant under G, contrary to the definition of G as a V^L-group. Thus F is a V^L-group and in particular G_i is a V^L-group. (If F be a maximal invariant subgroup of G, successive repetitions of the above argument show that any subgroup which is a member of any ordinary series of decomposition of G is a V^L-group.)

In G_i, let the maximal invariant L-subgroup be G_{i1}. G_{i1} is characteristic under G_i which is characteristic under G. Hence G_{i1} is characteristic under G. G_{i1} contains G_1, but as G_1 is the maximal invariant L-subgroup of G, $G_{i1} = G_1$. Thus the first two members of the V-series of G_i are identical respectively with the first two members of the V-series of G. Let Q_{i2} be the largest invariant K-subgroup in G_i/G_1. To this characteristic subgroup of G_i/G_1 there corresponds G_{i2}, a characteristic subgroup of G_i which is also characteristic under G. G_{i2} contains G_2, but since G_2 is the largest invariant subgroup of G containing G_1 in such wise that G_i/G_1 is a K-group, then $G_{i2} = G_2$ and the third member of the V-series of G_i is identical with the third member of the V-series of G. Repetition of the above argument proves that where $j \gtreqless i+1$, the j-th member of the V-series of G_i is identical with the j-th member of the V-series of $G_n = G$. The proof of Theorem I is thus complete.

THEOREM II. *If G of order LK is a $V^L{}_n$-group, then*

(a) *When m is even*

(1) *To $V^L{}_{m+i}/V^L{}_{m+1}$ corresponds a set of conjugate V^K-subgroups in $O(V^L{}_{m+1}/V^L{}_m$*
(2) *To $V^L{}_{m+2}/V^L{}_{m+1}$ " " " " " K-subgroups in $A(V^L{}_{m+1}/V^L{}_m$*

(b) *When m is odd*

(1) *To $V^L{}_{m+i}/V^L{}_{m+1}$ corresponds a set of conjugate V^L-subgroups in $O(V^L{}_{m+1}/V^L{}_m$*
(2) *To $V^L{}_{m+2}/V^L{}_{m+1}$ " " " " " L-subgroups in $A(V^L{}_{m+1}/V^L{}_m$*

Here the term "corresponds" is used in a familiar special sense to imply not only abstract isomorphism but an appropriate equivalence of automorphism; the exact nature of this relation will appear in the course of the proof.

Proof. If m is even, $V^L{}_{m+1}$ is the maximal invariant subgroup of G such that $V^L{}_{m+1}/V^L{}_m$ is an L-group. Hence $V^L{}_{m+2}/V^L{}_{m+1}$ is a K-group and $V^L{}_{m+i}/V^L{}_{m+1}$ contains no invariant L-subgroup and is thus a V^K-group. Similarly when m is odd $V^L{}_{m+2}/V^L{}_{m+1}$ is an L-group and $V^L{}_{m+i}/V^L{}_{m+1}$ is a V^L-group. To complete the proof of Theorem II it will now be shown that

When $m = 0$, and $V^L{}_m = V^L{}_0 =$ identity and $i > 1$, then
$V^L{}_i/V^L{}_1$ corresponds to a set of conjugate V^K-subgroups in $O(V^L{}_1)$ and
$V^L{}_2/V^L{}_1$ " " " " " " K-subgroups in $A(V^L{}_1)$.

Let J be the subgroup of $V^L{}_i$ which consists of all the operators in $V^L{}_i$ which are commutative with every operator in $V^L{}_1$. Since $V^L{}_1$ is a characteristic subgroup of $V^L{}_i$, both the central C of $V^L{}_1$ and the "centralizer" J of $V^L{}_1$ in $V^L{}_i$ are characteristic under $V^L{}_i$. The intersection of $V^L{}_i$ and J is C. Consider the quotient group J/C. Since $V^L{}_i$ is solvable, J/C is solvable, and the intersection D of some set of Sylow subgroups in J/C is characteristic under J/C. Obviously D is either an L-group or a K-group, if the order of J/C is > 1. Let E be the invariant subgroup of $V^L{}_i$ which corresponds to D in J/C. Suppose D to be an L-group. Then E would be an L-group since C (contained in $V^L{}_1$) is an L-group, and E would generate with $V^L{}_1$ an invariant L-subgroup of $V^L{}_i$ whose order would exceed that of $V^L{}_1$, contradicting the definition of $V^L{}_1$; hence D is not an L-group. Suppose D to be a K-group. Since every operator in E (a subgroup of J) is commutative with every operator in C (a subgroup of $V^L{}_1$), then E would be the direct product of C and F, the Sylow K-subgroup of E. Since E would be invariant under $V^L{}_i$, and F would be characteristic under E, F would be an invariant K-subgroup of $V^L{}_i$, contrary to the definition of $V^L{}_i$; hence D is not a K-group. Since D is neither a K-group nor an L-group, J/C is of order 1 and $J = C$. It follows that every operator of $V^L{}_i$ not in C sets up an automorphism in $V^L{}_1$.

Since $V^L{}_i/C$ is the group of all automorphisms set up in $V^L{}_1$ by $V^L{}_i$, then $V^L{}_i/C$ "corresponds" to a set of conjugate subgroups in $A(V^L{}_1)$. $V^L{}_1/C$ corresponds to $B(V^L{}_1)$ and hence $V^L{}_i/V^L{}_1$ corresponds to a set of conjugate subgroups in $A(V^L{}_1)/B(V^L{}_1) = O(V^L{}_1)$. Since $V^L{}_i/V^L{}_1$ is solvable and contains no invariant L-subgroups, the corresponding set of conjugate subgroups in $O(V^L{}_1)$ must be V^K-groups.

Since $V^L{}_2$ is solvable and the order of $V^L{}_1$ is prime to the index of $V^L{}_1$ under $V^L{}_2$, then $V^L{}_2$ contains exactly one set of conjugate K-subgroups simply isomorphic with $V^L{}_2/V^L{}_1$.[3] As no operator in such a K-subgroup is contained in C, every such operator sets up an automorphism in $V^L{}_1$. Hence such a K-subgroup is a subgroup of $A(V^L{}_1)$ and $V^L{}_2/V^L{}_1$ corresponds to this K-subgroup or to one of its conjugates under $A(V^L{}_1)$. This completes the proof of Theorem II.

The aspect of Theorems I and II which is of particular use in the listing of U^*-groups may be briefly stated as follows:

In any V series V_{i+2}/V_{i+1} is of order relatively prime to the order of V_{i+1}/V_i and corresponds to a set of conjugate subgroups in $A(V_{i+1}/V_i)$.

If (V_{i+1}/V_i) is cyclic and of odd order, $G = V_n$ is either V_{i+1} or V_{i+2}. If $A(V_{i+1}/V_i)$ contains no operator of order prime to the order of (V_{i+1}/V_i) then $G = V_n = V_{i+1}$. This is the case (for instance) where V_{i+1}/V_i is cyclic and of order 2^y. Theorem I indicates a method for determining all the V^L-groups stepped up from a given $V^L{}_1$. First determine all the $V^L{}_2$-groups stepped up from the given $V^L{}_1$; then determine all the $V^L{}_3$-groups stepped up from these $V^L{}_2$-groups; etc. That the entire list of V^L-groups thus obtained from a single $V^L{}_1$ comprises only a finite number of groups is a consequence of Theorem II.

The application of Theorems I and II where L is a power of a prime. If G is a solvable V^L-group of order LK it contains (according to Hall's theorem) a single set of conjugate subgroups of order K and can thus be expressed and in one way only as a transitive group of degree L. Where L is a power of a prime, G is thus a U-group as defined at the beginning of this paper, and its transitive expression of degree L is a U^*-group. In order to emphasize certain special properties of these groups it is convenient to handle them with a distinctive notation. Accordingly, when L is a power of a prime, V will be replaced by U. The symbol $V^L{}_i$ becomes $U_i(p^a)$ and the U-series of $U_n(p^a)$ may be written,

$$1 = U_0(p^a) < U_1(p^a) \cdots < U_{n-1}(p^a) < U_n(p^a) = G.$$

In the sense of this notation, a solvable group G is a U-group if some one prime factor of g divides the order of every proper invariant subgroup. As previously shown, no solvable group can have this relation to more than one

[3] P. Hall, *Journal of the London Mathematical Society*, vol. 3 (1932), p. 98.

prime. The equation $G = U_n(p^a)$ means that G is a U-group with respect to the prime p and that it has $n-1$ proper subgroups in its U-series. If the fuller notation $U_n(p^a, K)$ is used, it will be understood in addition that G is precisely of order p^aK with K prime to p.

According to this notation if $K = 1$, U is a prime power group U_1, and U^* is its regular expression. Where $a = 1$, U^* is either the metacyclic group of degree p and order $p(p-1)$ or one of its invariant subgroups, since no other solvable transitive groups of prime degree exist. Both of these categories, the prime power groups and the metacyclic groups with their invariant subgroups are too well known to require discussion here.

In $U_n(p^a, K)$ the U-series determines a sequence of integers $a_1 \cdots a_f$ such that $\sum_1^f a_i = a$ and p^{a_i} is the index of U_{2i-2} under U_{2i-1}. If n is even, $f = n/2$; if n is odd $f = (n+1)/2$. This sequence of integers will be called the U-sequence of U. Hall's results [4] show that if F is a group of order p^y, the order of the p-Sylow subgroup in $A(F)$, and *a fortiori* the order of the p-Sylow subgroup in $O(F)$ does not exceed $p^{y(y-1)/2}$. But U_{2i-1}/U_{2i-2} is a p-group and U_n/U_{2i-1} corresponds to a set of conjugate subgroups in $O(U_{2i-1}/U_{2i-2})$. Hence the U-sequence of U_n is subject to the restriction that $\sum_{i+1}^f a_j \leqq a_i(a_i - 1)/2$. It follows that the sequence $a_1 \cdots a_f$ dominates ($\geqq$) the arithmetic sequence $f \cdots 2, 1$. Hence if the number t is in the U-sequence, it is one of the last t members. However, it seems improbable that in the case of every sequence satisfying the above restriction, there exists a group for which it is the U-sequence.

The problem of constructing all the groups "stepped up" from a given U_1 may be attacked in various ways. The methods outlined below have proven convenient in practice and adequate for U^*-groups of degrees < 64. They are based primarily on the consideration that

(a) *Where i is even (U_{i+2}/U_{i+1}) corresponds to a set of conjugate K-subgroups in $A(U_{i+1}/U_i)$.*

(b) *Where i is odd (U_{i+2}/U_{i+1}) corresponds to a set of conjugate p^w-subgroups in $A(U_{i+1}/U_i)$.*

The construction of all possible U_2 groups stepped up from a given U_1 is easy since there is exactly one such group for every set of conjugate K-sub-

[4] P. Hall, *Proceedings of the London Mathematical Society*, ser. 2, vol. 36 (1933-34), pp. 36, 37.

groups in $A(U_1)$. U^*_2 is thus always a subgroup of the holomorph of U^*_1. The search for all the U_2 groups stepped up from all possible U_1 groups of given order is much facilitated by a theorem of Hall. His results [5] show that if F is a prime power group and E the elementary group of the same order as F, then every K-subgroup in $A(F)$ is simply isomorphic with a K-subgroup in $A(E)$.

The problem of constructing all possible U_3 groups stepped up from a given U_2 is somewhat more intricate. Let M_2 be a maximal K-subgroup of U_2. M_2 is simply isomorphic to U_2/U_1, and with U_1 generates U_2. Each M_2 is also a maximal K-subgroup of U_3 and contained in its invariant subgroup U_2. If N_2 and N_3 are the normalizers of M_2 in U_2 and U_3 respectively then N_3/N_2 is simply isomorphic with U_3/U_2. Since U_3/U_2 corresponds to a set of conjugate p-subgroups in $A(U_2/U_1)$ [and therefore in $A(M_2)$], U_3 may be generated by adjoining to U_2 generators $j_1 \cdots j_s$ which

(a) generate with U_1 a p-Sylow subgroup of U_3,

(b) transform the operators of M_2 according to the group (U_3/U_2).

If h_i is the lowest power of j_i contained in U_1, then h_i must set up the identity automorphism in M_2. Thus h_i is an operator in H_2, the subgroup of U_1 which consists of all those operators in U_1 which are commutative with every operator in M_2. Let C_i be the central of U_i. Since every operator in U_i which is not in C_1 sets up an automorphism in U_1, $C_1 \gtreqless C_i$. Thus H_2 always contains C_2 and it is convenient to distinguish two categories of cases (1) $H_2 = C_2$, (2) $H_2 > C_2$.

(1) $H_2 = C_2$. This case always arises when U_1 is abelian, since under these circumstances, any operator in U_1 commutative with every operator in M_2 is in C_2. But there are numerous cases where $H_2 = C_2$ although U_1 is not abelian. In all cases of this category however, h_i sets up the identity automorphism in U_1 and U_3/U_1 corresponds to a set of conjugate V^K-subgroups in $A(U_1)$. It is convenient to divide this category of U_2 groups according to whether (a) $H_2 = C_2 = 1$ or (b) $H_2 = C_2 > 1$.

(a) $H_2 = C_2 = 1$. Here M_2 and $j_1 \cdots j_s$ generate a V^K-subgroup of $A(U_1)$ simply isomorphic with U_3/U_1 and there is a one to one correspondence between the U_3 groups derived from a given U_2 and the sets of conjugate V^K-subgroups in $A(U_1)$ containing the given M_2 (or one of its conjugates) as an invariant subgroup of index p^v. Every such U_3 group may be expressed not only as a U^*-group but as a subgroup of the holomorph of its U^*_1.

---

[5] P. Hall, *Proceedings of the London Mathematical Society*, ser. 2, vol. 36 (1933-34), pp. 37, 38.

(b) $H_2 = C_2 > 1$. In this case, a variety of abstract groups U_3 is sometimes obtainable from the same U_2 and the same subgroup of automorphisms of U_1. To choose the j_i's so that each $h_i = 1$ is always possible; if they are so chosen, the resulting U_3 may be written as a subgroup of the holomorph of U^*_1 exactly as in the case where $H_2 = C_2 = 1$. The generators so chosen will be called j'_i's, and the U_3-group thus generated will be called a principal U_3-group. But where $H_2 = C_2 > 1$, it is frequently possible to choose the j_i's so that they set up respectively in U_2 the same automorphisms as the j'_i's but so that not every h_i is identity. With generators so chosen the resulting U_3 is sometimes (not always) distinct from the principal U_3 group which corresponds to the same set of conjugate V^K-subgroups in $A(U_1)$. Such a distinct U_3-group will be called a collateral U_3-group; it cannot be written as a subgroup of the holomorph of its U^*_1.

It is thus shown that when $H_2 = C_2 > 1$, the correspondence between U_3-groups and the suitable sets of V^K-subgroups in $A(U_1)$ is not always one to one but in certain cases many to one. In an extended program for the construction of U_3 groups it would be desirable to develop rules for the number of collateral U_3 groups belonging to any principal U_3. But the subject is an intricate one and, within the field covered by this paper, the cases where collateral groups arise occasion no great difficulty and have been worked out individually.

(2) $H_2 > C_2$. In this case U_3/U_1 may correspond to a set of conjugate V^K-subgroups not in $A(U_1)$ but in $O(U_1)$. In a general program, the U_3 groups stepped up from U_2 groups of this category would require detailed treatment, but inspection of all U_2 groups of low order has shown that the relation $H_2 > C_2$ is impossible if U_2 is of order < 96. A U_3 group stepped up from this sort of U_2 would have an order $\gtreqless 192$ and a degree $\gtreqless 64$. Further discussion of this case is therefore omitted.

Considerations similar to those already given hold for the construction of U_n groups where $n > 3$. But sufficient analysis has been carried out to show that, for degrees < 64, there is only one instance of a U^*_n where $n > 3$; so that no further discussion of the general theory here required is now attempted.

The U^* groups of degrees < 64 and their corresponding U groups. If U^* is of order p^aK and degree $p^a < 64$ *then* $a < 6$ and the presumptive U-sequences are as follows:

a	*U-sequences*		
1	1		
2	2		
3	3	2, 1	
4	4	3, 1	
5	5	4, 1	3, 2.

But consider the sequence 3, 2. Here p must be 2 and U_1 of order 8. $O(U_1)$ must contain a set of conjugate solvable subgroups of order $4K$ (where $K > 1$) which contain no invariant subgroup of order 2 or 4. If G is of order 8, no $O(G)$ meets this condition and so, where U^* is of degree < 64, the U-sequence 3, 2 does not occur.

If the cases where $a = 1$ or $K = 1$ are omitted (for reasons already explained), and if it is borne in mind that, when U_{i+1}/U_i is of order 2, then $U_{i+1} = U_n$, the following table gives the possibilities for a U^*-group of degree < 64.

a	*U-sequence*	p	n *in* U_n
2	2	2 or 3 or 5 or 7	2
3	3	2 or 3	2
	2, 1	2	3
		3	3 or 4
4	4	2	2
	3, 1	2	3
5	5	2	2
	4, 1	2	3.

From the table it appears that, where $p > 2$, n exceeds 2 only where U^* is of degree 27 and U_1 of order 9 and therefore abelian. U_3/U_1 and U_4/U_1 in this case must correspond to subgroups of $A(U_1)$ which are of order $3K$ but contain no invariant subgroup of order 3. Hence U_1 is not cyclic. If U_1 is not cyclic and of order 9, the only subgroups of $A(U_1)$ which have orders divisible by 3 but contain no invariant subgroups of order 3 are $A(U_1)$ itself and its

single subgroup of index 2. If U_n/U_1 corresponds to either of these groups, U_2 is such that $H_2 = C_2 = 1$ and the correspondence between the U_n-groups and the subgroups of $A(U_1)$ is one to one. Thus, if $p > 2$ and $n > 2$, there are only two U^*-groups of degree < 64. Of the corresponding U-groups one ($n = 4$) may be written as the holomorph (of order 432) of the non-cyclic group of order 9; the other ($n = 3$) may be written as the positive half of this holomorph.

Where $p = 2$, n never exceeds 3 and attains this figure only for the sequences 2, 1 or 3, 1 or 4, 1.

(a) If U_1 is of order 4, $K = 3$, and U_3 is of order 24 and a principal U_3 group.

(b) If U_1 is of order 8, $K = 3$, or 7 or 21 and U_3 would be of order 48, 112 or 336. But in the latter two cases U_3/U_1 would correspond to a set of conjugate subgroups of order 14 or 42 in $A(U_1)$ where U_1 is the elementary group of order 8. As this $A(U_1)$ contains no subgroups of orders 14 or 42, U_3 is of order 48, where U_1 is of order 8. Since, in the U_2 groups of order 24, $H_2 = C_2 > 1$, the possibility of collateral groups of order 48 must be considered.

(c) If U_1 is of order 16, $K = 3, 5, 7, 9, 15$, or 21. If $K > 3$, U_1 is the elementary group of order 16 and U_3/U_1 corresponds to a set of conjugate subgroups of order $2K$ in $A(U_1)$. As this $A(U_1)$ contains no subgroups of order 14 or 42 there are no U_3 groups of order 224 or 672. Where $K = 5$, 9 or 15, U_2 is such that $H_2 = C_2 = 1$ and no collateral U_3 groups of orders 160, 288 or 480 are possible. Where $K = 3$, U_1 is not necessarily elementary and U_2 is sometimes such that $H_2 = C_2 > 1$, so that collateral groups of order 96 are to be looked for.

It is thus shown that if U^*_n is of degree < 64, $n > 2$ only for the orders 24, 48, 96, 160, 216, 288, 432 and 480, and that, for orders other than 48 and 96, collateral U_3 groups are impossible. A U_4 group occurs only in the order 432. As all the groups of orders 48 and 96 are known, a discussion of the number of collateral groups is superfluous and the list of groups of these orders is taken from the published tables. In all other orders the correspondence between U_n groups ($n > 1$) and the suitable sets of conjugate subgroups in $A(U_1)$ is one to one.

The following table summarizes the U^*-groups of degree < 64 where $a > 1$ and $K > 1$

Degree	K	Orders of U_2	U_3	U_4
4	3	12	..	...
8	Factors of 21	$8K$	24	...
9	Factors of 16	$9K$	..	...
16	3, 5, 7, 9, 15, 21	$16K$	48	...
25	Factors of 96	$25K$	..	...
27	2, 4, 8, 13, 16, 26, 32	$27K$	216	432
32	3, 5, 7, 9, 15, 21, 31, 63, 155	$32K$	96, 160 288, 480	...
49	Proper factors of 288	$49K$	..	...

As appears from the following list there are just 212 of these groups. To these are added the 101 U^*-groups of degree < 64 where $a = 1$, and the 83 U^*_1-groups of degree < 64 where $a > 1$. Thus the basis groups whose U^* expressions are of degree < 64 are 396 in number. Any solvable group which contains no Sylow subgroup of order $\gtreqless 64$ may be constructed from these 396 groups. The number of such groups is hard to estimate, but, if the basis is limited to the 102 U^*_1-groups in the above list, the calculation is easy and shows that approximately 1.75×10^8 groups can be constructed from this portion of the basis groups. There can therefore be little doubt that the number of groups which can be constructed from all the 396 basis groups of degree < 64 easily exceeds 10^{10}.

The Non-regular U^*-groups of degrees 4, 8, 9, 16, 25, 27, 32 *and* 49. The authors have proven the completeness of the following list—e. g. they have shown that where $p = 5$ and G is the non-cyclic group of order p^2, $A(G)$ contains exactly 29 sets of conjugate solvable K-subgroups. But these arguments are too lengthy for inclusion here.

The generators of the various U^*-groups are given in a particular order. The symbol X_x stands for a generator, and

$$A_1 \cdots A_a \mid B_1 \cdots B_b \mid C_1 \cdots M_m \mid N_1 \cdots N_n$$

means that the generators $A_1 \cdots A_a$ generate U_1; the generators $B_1 \cdots B_b$ extend U_1 to U_2; $C_1 \cdots C_c$ extend U_2 to U_3; etc. Furthermore the generators are arranged in such order that the group generated by the first x generators is invariant under the group generated by the first $x + 1$ generators.

In the degrees 16 and 32, each principal U_3 group of order 48 or 96 respectively, if it has collateral groups, is labelled '*pr*' and bracketed with its collateral groups labelled '*co.*'

DEGREE 4

A_1 = ac.bd
A_2 = ab.cd
B_1 = cbd

U*			Order
#1	A_1A_2	B_1	12

DEGREE 8

A_1 = abcd.efgh
A_2 = aecg.bhdf
A_3 = ac.bd.eg.fh
A_4 = abcd.ef.gh
A_5 = ae.bf.cg.dh
B_1 = bef.dgh
B_2 = ecbgdhf
B_3 = cbd.gfh
C_1 = ae.bh.cg.df

U*				Order
#1	A_1A_2	B_1		24
#2	$A_3A_4A_5$	B_3		24
#3	″	B_2		56
#4	″	B_2B_3		168
#5	A_3A_4	B_3	C_1	24

DEGREE 9

A_1 = adhbeicfg
A_2 = abc.def.ghi
A_3 = adg.beh.cfi
B_1 = bc.dg.ei.fh
B_2 = dg.eh.fi
B_3 = bdcg.efih
B_4 = beci.dhgf
B_5 = bhdecfgi

U*			Order
#1	A_1	B_1	18
#2	A_2A_3	B_1	18
#3	″	B_2	18
#4	″	B_1B_2	36
#5	″	B_3	36
#6	″	B_3B_2	72
#7	″	B_3B_4	72
#8	″	B_5	72
#9	″	B_5B_2	44

DEGREE 16

A_1 = ac.bd.eg.fh.ik.jl.mo.np
A_2 = ab.cd.ef.gh.ij.kl.mn.op
A_3 = ae.bf.cg.dh.im.jn.ko.lp
A_4 = ai.bj.ck.dl.em.fn.go.hp
A_5 = abcd.efgh.ijkl.mnop
A_6 = aeim.bfjn.cgko.dhlp
A_7 = aecg.bhdf.imko.jpln
A_8 = aick.bjdl.emgo.fnhp
B_1 = ecbgdhf.mkjolpn
B_2 = cbd.gfh.kjl.onp
B_3 = cbd.eim.flo.gjp.hkn
B_4 = bef.dgh.jmn.lop
B_5 = bep.cik.dmf.ghl.jon
B_6 = cplkm.bgone.djfhi
C_1 = ai.bl.ck.dj.em.fp.go.hn
C_2 = aiem.blfp.ckgo.djhn
C_3 = ai.bl.ck.dj.ep.fo.gn.hm
C_4 = aick.bldj.epgn.fohm

U*					Order
# 1	$A_1A_2A_6$	B_2			48
# 2	$A_5A_7A_4$	B_4			48
# 3	$A_5A_7A_8$	B_4			48
# 4	A_5A_6	B_5			48
# 5	$A_1A_2A_3A_4$	B_2			48
# 6	″	B_3			48
# 7	″	B_6			80
# 8	″	B_1			112
# 9	″	B_3B_2			144
#10	″	B_6B_1			240
#11	″	B_1B_2			336
#12	$A_1A_2A_3$	B_2	C_1	pr	48
#13	″	″	C_2	co	48
#14	A_5A_7	B_4	C_3	pr	48
#15	″	″	C_4	co	48

DEGREE 25

A_1 = aforvbgkswchltxdımpyejnqu
A_2 = abcde.fghij.klmno.pqrst.uvwxy
A_3 = afkpu.bglqv.chmrw.dinsx.ejoty
B_1 = be.cd.fu.gy.hx.iw.jv.kp.lt.ms.nr.oq
B_2 = fu.gv.hw.ix.jy.kp.lq.mr.ns.ot
B_3 = bwh.cto.dlq.eix.fmn.gjp.kyv.rus
B_4 = bxg.cqm.ehy.dos.fir.ptv.uwn.klj
B_5 = bced.fkup.gmys.hoxq.ilwt.jnvr
B_6 = bced.fpuk.xlht.qwoi.gryn.jsvm
B_7 = bfeu.ckdp.xqho.litw.gjyv.rmns
B_8 = bced.ghji.lmon.qrts.vwyx
B_9 = fu.kp.bced.gwjx.hyiv.lros.mtnq
B_{10} = bpcfekdu.hjonxvqr.iylswgtm
B_{11} = bkdfepcu.hyqmxgos.ijtrwvln

U*			ORDER	U*			ORDER
#1	A_1	B_1	50	#17	A_2A_3	$B_{[illegible]}$	200
#2	"	B_5	100	#18	"	$B_3B_1B_2$	300
#3	A_2A_3	B_1	50	#19	"	B_3B_5	300
#4	"	B_2	50	#20	"	B_3B_6	300
#5	"	B_3	75	#21	"	$B_{10}B_2$	400
#6	"	B_1B_2	100	#22	"	$B_6B_7B_5$	400
#7	"	B_5	100	#23	"	B_5B_8	400
#8	"	B_6	100	#24	"	B_3B_{10}	600
#9	"	B_8	100	#25	"	B_3B_{11}	600
#10	"	B_9	100	#26	"	$B_3B_5B_2$	600
#11	"	B_3B_1	150	#27	"	$B_6B_7B_4$	600
#12	"	B_3B_2	150	#28	"	$B_{10}B_2B_8$	800
#13	"	B_4B_2	200	#29	"	$B_3B_{10}B_2$	1200
#14	"	B_5B_2	200	#30	"	$B_6B_7B_4B_5$	1200
#15	"	B_7B_2	200	#31	"	$B_6B_7B_4B_5B_{10}$	2400
#16	"	B_6B_7	200				

DEGREE 27

A_1 = ajvdmzhqtbkwenαiruclxfoygps
A_2 = adhbeıcfg.jmqknrlop.svztwαuxy
A_3 = abc.def.ghi.jkl.mno.pqr.stu.vwx.yzα
A_4 = adg.beh.cfı.jmp.knq.lor.svy.twz.uxα
A_5 = ajs.bkt.clu.dmv.enw.fox.gpy.hqz.irα
A_6 = ajs.bkt.clu.dnx.eov.fmw.grz.hpα.iqy
B_1 = bc.dg.ei.fh.js.ku.lt.my.nα.oz.pv.qx.rw
B_2 = bc.dg.eı.fh.kl.mp.nr.oq.tu.vy.wα.xz
B_3 = js.kt.lu.mv.nw.ox.py.qz.rα
B_4 = dg.eh.fi.js.kt.lu.my.nz.oα.pv.qw.rx
B_5 = dg.eh.fı.mp.nq.or.vy.wz.xα
B_6 = bc.ef.hi.js.ku.lt.mv.nx.ow.py.qα.rz
B_7 = bdcg.efih.kmlp.norq.tvuy.wxαz
B_8 = beci.dhgf.knlr.mqpo.twuα.vzyx
B_9 = js.bdcg.efih.kvly.mupt.nxrz.oαqw
B_{10} = js.beci.dhgf.kwlα.mzpx.tnur.vqyo
B_{11} = djqs.oqαw.ekht.mryx.fliu.npzv
B_{12} = dogα.jwsq.emhy.kxtr.fnız.lvup
B_{13} = beghcidf.kupqlrmo.twyzuαvx
B_{14} = js.beghcidf.kwpzlαmx.tnyqurvo
B_{15} = bc.dwjogqsα.evknhptz.fxlmiruy
B_{16} = bdjemnqzufpwl.cgsiyαxokhvrt
C_1 = ajs.bkt.clu.dnx.eov.fmw.grz.hpα.ıqy
D_1 = dg.eh.fi.js.kt.lu.my.nz.oα.pv.qw.rx

U*			ORDER	U*					ORDER
#1	A_1	B_1	54	#23	$A_3A_4A_5$	B_{13}			216
#2	A_2A_5	B_1	54	#24	"	B_{14}			216
#3	"	B_2	54	#25	"	B_7B_3			216
#4	"	B_3	54	#26	"	B_7B_5			216
#5	"	B_2B_3	108	#27	"	B_7B_4			216
#6	A_2A_6	B_2	54	#28	"	B_9B_5			216
#7	$A_3A_4A_6$	B_2	54	#29	"	B_7B_8			216
#8	"	B_4	54	#30	"	$B_7B_{[illegible]}$			216
#9	"	B_2B_4	108	#31	"	$B_2B_3B_5$			216
#10	"	B_{11}	108	#32	"	B_{16}			351
#11	"	B_{15}	216	#33	"	$B_{13}B_3$			432
#12	"	$B_{11}B_6$	216	#34	"	$B_7B_5B_3$			432
#13	"	$B_{11}B_{12}$	216	#35	"	$B_7B_8B_3$			432
#14	"	$B_{15}B_6$	432	#36	"	$B_{13}B_5$			432
#15	$A_3A_4A_5$	B_1	54	#37	"	$B_{13}B_4$			432
#16	"	B_2	54	#38	"	$B_{14}B_5$			432
#17	"	B_3	54	#39	"	$B_{14}B_4$			432
#18	"	B_2B_3	108	#40	"	$B_{16}B_1$			702
#19	"	B_2B_4	108	#41	"	$B_{13}B_5B_3$			864
#20	"	B_2B_5	108	#42	A_3A_4	B_7B_8	C_1		216
#21	"	B_7	108	#43	A_3A_4	B_7B_8	C_1	D_1	432
#22	"	B_9	108						

DEGREE 32

A_1=ac.bd.eg.fh.ik.jl.mo.np.AC.BD.EG.FH.IK.JL.MO.NP
A_2=ab.cd.ef.gh.ij.kl.mn.op.AB.CD.EF.GH.IJ.KL.MN.OP
A_3=ae.bf.cg.dh.im.jn.ko.lp.AE.BF.CG.DH.IM.JN.KO.LP
A_4=ai.bj.ck.dl.em.fn.go.hp.AI.BJ.CK.DL.EM.FN.GO.HP
A_5=aA.bB.cC.dD.eE.fF.gG.hH.iI.jJ.kK.lL.mM.nN.oO.pP
A_6=aeim.bfjn.cgko.dhlp.AEIM.BFJN.CGKO.DHLP
A_7=aAeEiImM.bBfFjJnN.cCgGkKoO.dDhHlLpP
A_8=aA.bB.cC.dD.eM.fN.gO.hP.iI.jJ.kK.lL.mE.nF.oG.pH
A_9=abcd.efgh.ijkl.mnop.ABCD.EFGH.IJKL.MNOP
A_{10}=aecg.bhdf.imko.jpln.AECG.BHDF.IMKO.JPLN
A_{11}=aick.bjdl.emgo.fnhp.AICK.BJDL.EMGO.FNHP
A_{12}=aAiIcCkK.bBjJdDlL.eEmMgGoO.fFnNhHpP
A_{13}=aAiI.bBjJ.cCkK.dDlL.eEmM.fFnN.gGoO.hHpP
A_{14}=aAcC.bBdD.eEgG.fFhH.iKkI.jLlJ.mOoM.nPpN
A_{15}=aA.bB.cC.dD.eE.fF.gG.hH.iK.jL.kI.lJ.mO.nP.oM.pN
A_{16}=aA.bD.cC.dB.eM.fP.gO.hN.iI.jL.kK.lJ.mE.nH.oG.pF
A_{17}=aA.bJ.cC.dL.eO.fH.gM.hF.iI.jB.kK.lD.mG.nP.oE.pN
A_{18}=aA.bI.cE.dM.eC.fK.gG.hO.iB.jJ.kF.lN.mD.nL.oH.pP
A_{19}=aAiI.bFjN.cCkK.dHlP.eEmM.fBnJ.gGoO.hDpL

B_1=cbd.gfh.kjl.onp.CBD.GFH.KJL.ONP
B_2=AcbCdDB.EgfGhHF.IkjKlLJ.MonOpPN
B_3=cbd.eim.gjp.hkn.flo.CBD.EIM.GJP.HKN.FLO
B_4=cplkm.bgone.djfhi.CPLKM.BGONE.DJFHI
B_5=eim.fjn.gko.hlp.EIM.FJN.GKO.HLP
B_6=AbgLoGKinEOHMDdcePBhNFImCfJkjlp
B_7=bef.dgh.jmn.lop.BEF.DGH.JMN.LOP
B_8=iAI.jBJ.kCK.lDL.mEM.nFN.oGO.pHP
B_9=bef.dgh.iAI.jEN.kCK.lGP.mFJ.nBM.oHL.pDO
B_{10}=jmnAK.lopCI.bPkBf.dNiDh.eMFLG.gOHJE
B_{11}=bep.cik.dmf.ghl.jon.BEP.CIK.DMF.GHL.JON
B_{12}=bAP.cio.dIB.fEL.gmk.hMF.jON.lGD.nKJ.pCH

C_1=aA.bD.cC.dB.eE.fH.gG.hF.iI.jL.kK.lJ.mM.nP.oO.pN
C_2=aAeEiImM.bDfHjLnP.cCgGkKoO.dBhFlJpN
C_3=aA.bD.cC.dB.eM.fP.gO.hN.iI.jL.kK.lJ.mE.nH.oG.pF
C_4=aAiI.bDjL.cCkK.dBlJ.eMmE.fPnH.gOoG.hNpF
C_5=aAiI.bDjL.cCkK.dBlJ.eEmM.fHnP.gGoO.hFpN
C_6=aA.bB.cC.dD.eG.fH.gE.hF.iJ.jI.kL.lK.mP.nO.oN.pM
C_7=aA.bI.cE.dM.eC.fK.gG.hO.iB.jJ.kF.lN.mD.nL.oH.pP
C_8=aA.bD.cC.dB.eH.fG.gF.hE.iI.jL.kK.lJ.mP.nO.oN.pM
C_9=aAiIcCkK.bDjLdBlJ.eHmPgFoN.fGnOhEpM
C_{10}=aA.bD.cC.dB.eH.fG.gF.hE.iK.jJ.kI.lL.mN.nM.oP.pO
C_{11}=aAcC.bDdB.eHgF.fGhE.iKkI.jJlL.mNoP.nMpO
C_{12}=aAcC.bDdB.eHgF.fGhE.iIkK.jLlJ.mPoN.nOpM
C_{13}=aAiI.bDjL.cCkK.dBlJ.eHmP.fGnO.gFoN.hEpM
C_{14}=aA.bB.cC.dD.eP.fM.gN.hO.iK.jL.kI.lJ.mF.nG.oH.pE

U*					ORDER
#1	$A_1 A_2 A_7$	B_1			96
#2	$A_1 A_2 A_6 A_5$	B_1			96
#3	"	B_2			224
#4	"	$B_2 B_1$			672
#5	$A_1 A_2 A_6 A_8$	B_1			96
#6	$A_1 A_2 A_3 A_4 A_5$	B_1			96
#7	"	B_3			96
#8	"	B_4			160
#9	"	B_2			224
#10	"	$B_1 B_5$			288
#11	"	$B_3 B_4$			480
#12	"	$B_2 B_1$			672
#13	"	$B_2 B_5$			672
#14	"	$B_2 B_3$			672
#15	"	B_6			992
#16	"	$B_2 B_1 B_5$			2016
#17	"	$B_6 B_4$			4960
#18	$A_9 A_{10} A_{12}$	B_7			96
#19	$A_9 A_{10} A_{11} A_5$	B_7			96
#20	$A_9 A_{10} A_{13}$	B_7			96
#21	$A_9 A_{10} A_4 A_5$	B_7			96
#22	"	B_8			96
#23	"	B_9			96
#24	"	$B_7 B_9$			288
#25	$A_9 A_{10} A_{11} A_{14}$	B_7			96
#26	"	B_9			96
#27	"	$B_7 B_9$			288
#28	$A_9 A_{10} A_{12} A_{15}$	B_7			96
#29	"	B_{10}			160
#30	$A_9 A_6 A_5$	B_{11}			96
#31	$A_9 A_6 A_{16}$	B_{11}			96
#32	$A_9 A_6 A_{17}$	B_{11}			96
#33	$A_1 A_2 A_3 A_4 A_{18}$	B_3			96
#34	$A_3 A_9 A_{19}$	B_{12}			96
#35	$A_1 A_2 A_6$	B_1	C_1	pr	96
#36	"	"	C_2	co	96
#37	"	"	C_3	pr	96
#38	"	"	C_4	co	96
#39	$A_1 A_2 A_3 A_4$	B_1	C_1	pr	96
#40	"	"	C_5	co	96
#41	"	"	C_3		96
#42	"	B_3	C_3		96
#43	"	B_4	C_6		160
#44	"	$B_1 B_5$	C_1		288
#45	"	"	C_3		288
#46	"	"	C_7		288
#47	"	$B_4 B_3$	C_6		480
#48	$A_9 A_{10} A_{11}$	B_7	C_8	pr	96
#49	"	"	C_9	co	96
#50	"	"	C_{10}	pr	96
#51	"	"	C_{11}	co	96
#52	$A_9 A_{10} A_4$	B_7	C_8	pr	96
#53	"	"	C_{12}	co	96
#54	"	"	C_{13}	co	96
#55	"	"	C_{13}		96
#56	$A_9 A_6$	B_{11}	C_{14}		96

DEGREE 49

A_1 = ahrDMTb′bisENUc′cjtFHOd′dkuGIPe′eloAJQf′fmpBKRg′gnqCLSa′
A_2 = abcdefg.hijklmn.opqrstu.ABCDEFG.HIJKLMN.OPQRSTU.a′b′c′d′e′f′g′
A_3 = ahoAHOa′.bipBIPb′.cjqCJQc′.dkrDKRd′.elsELSe′.fmtFMTf′.gnuGNUg′

B_1 = bg.cf.de.ha′.ig′.jf′.ke′.ld′.mc′.nb′.oO.pU.qT.rS.sR.tQ.uP.AH.BN.CM.DL.EK.FJ.GI
B_2 = hb′.ic′.jd′.ke′.lf′.mg′.na′.oQ.qS.sU.uP.pR.rT.tO.AK.DN.GJ.CM.FI.BL.EH
B_3 = he′.if′.jg′.ka′.lb′.mc′.nd′.sT.uO.pQ.rS.tU.oP.qR.AM.DI.GL.CH.FK.BN.EJ
B_4 = bc′.eB.cS.uL.gm.dN.fr.PD.iM.e′T.g′C.kq.oG.b′n.OI.EK.a′s.Rh.Hf′.jA.tQ
B_5 = bce.gfd.hoH.iqL.jsI.kuM.lpJ.mrN.ntK.Aa′O.Bc′S.Ce′P.Dg′T.Eb′Q.Fd′U.Gf′R
B_6 = bec.gdf.uLl.PDd′.ipM.Fe′T.g′UC.Jkq.c′SB.Gb′O.mrN.Ino.Ea′R.Hjt.Khs.Af′Q
B_7 = hlt.iJu.jKo.kLp.lMq.mNr.nHs.ARb′.BSc′.CTd′.DUe′.EOf′.FPg′.GQa′
B_8 = bec.dfg.hnj.ikl.ots.upq.AEG.CFD.IHK.JLM.OQR.PUT.b′f′a′.d′g′e′
B_9 = bugP.iFg′J.ekde′.LUDp.cMfC.qd′Tl.c′GmI.EHKA.Sf′rj.b′hna′.BRNs.QotO
B_{10} = bigg′.uJPF.eLdD.kpe′U.cqfT.MlCd′.c′EmK.GAIH.Sb′rn.f′a′jh.BQNt.ROso
B_{11} = bc′gm.cSfr.dNeB.hTa′q.iHg′A.jCf′M.kse′R.lnd′b′.oDOL.ptUQ.uIPG.EJKF
B_{12} = bEPAgKuH.iGJc′g′IFm.cb′Ca′fnMh.qf′lSTjd′r.eQe′Odtko.LRpBDsUN
B_{13} = biEGPJAc′gg′KIuFHm.cqb′f′Cla′SfTnjMd′hr.eLQRe′pOBdDtskUoN

U^*			ORDER
#1	A_1	B_1	98
#2	·	B_5	147
#3	·	$B_1 B_5$	294
#4	A_2A_3	B_1	98
#5	·	B_2	98
#6	·	B_5	147
#7	·	B_6	147
#8	·	B_7	147
#9	·	$B_1 B_2$	196
#10	·	B_9	196
#11	·	$B_5 B_1$	294
#12	·	$B_6 B_1$	294
#13	·	$B_5 B_2$	294
#14	·	$B_6 B_3$	294
#15	·	$B_7 B_1$	294
#16	·	$B_7 B_3$	294
#17	·	$B_8 B_3$	294
#18	·	$B_6 B_4$	294
#19	·	$B_9 B_2$	392
#20	·	$B_9 B_{10}$	392
#21	·	B_{12}	392
#22	·	$B_5 B_6$	441
#23	·	$B_5 B_9$	588
#24	·	$B_5 B_1 B_3$	588
#25	·	$B_6 B_1 B_3$	588
#26	·	$B_7 B_1 B_3$	588
#27	$A_2 A_3$	$B_6 B_4 B_1$	588
#28	″	$B_6 B_{11}$	588
#29	″	B_{13}	784
#30	″	$B_{12} B_2$	784
#31	″	$B_{12} B_{10}$	784
#32	″	$B_5 B_6 B_1$	882
#33	″	$B_5 B_6 B_3$	882
#34	″	$B_5 B_6 B_4$	882
#35	″	$B_5 B_{12}$	1176
#36	″	$B_5 B_9 B_2$	1176
#37	″	$B_5 B_9 B_{10}$	1176
#38	″	$B_6 B_1 B_3 B_4$	1176
#39	″	$B_9 B_{10} B_6$	1176
#40	″	$B_9 B_{10} B_7$	1176
#41	″	$B_{13} B_2$	1568
#42	″	$B_5 B_6 B_1 B_3$	1764
#43	″	$B_5 B_6 B_1 B_4$	1764
#44	″	$B_5 B_6 B_{11}$	1764
#45	″	$B_5 B_{13}$	2352
#46	″	$B_5 B_{12} B_2$	2352
#47	″	$B_5 B_{12} B_{10}$	2352
#48	″	$B_9 B_{10} B_6 B_{11}$	2352
#49	″	$B_5 B_6 B_1 B_3 B_4$	3528
#50	″	$B_5 B_9 B_{10} B_6$	3528
#51	″	$B_5 B_{13} B_2$	4704
#52	″	$B_5 B_9 B_{10} B_6 B_{11}$	7056

UNIVERSITY OF CHICAGO.

ARC-PRESERVING TRANSFORMATIONS.

By G. T. Whyburn.

1. *Introduction.* If A is a compact continuum, a single valued transformation $T(A) = B$, where A and B are contained in a metric space, will be said to be *arc preserving* provided that the image of every simple arc in A is either a simple arc or a single point in B. Similarly, T will be said to be *irreducible* provided that no proper subcontinuum of A maps onto all of B under T.

Clearly any topological transformation is both arc preserving and irreducible. Obviously, also, if A contains no arc, then any transformation defined on A is arc preserving; and if B is itself a simple arc, then any continuous transformation sending A into B is arc preserving. The function $y = x^3 - x$ transforms the interval $(-2, 2)$ of the X-axis continuously into the interval $(-6, 6)$ of the Y-axis; and in this case the transformation is both arc preserving and irreducible; but is not a homeomorphism, since the roots of $x^3 - x - y$ are all real for $|y| \leq 2\sqrt{3}/9$ and are distinct for $|y| < 2\sqrt{3}/9$ whereas there is only one real root for $|y| > 2\sqrt{3}/9$.

Thus we see that even in case A and B are simple arcs (and therefore homeomorphic sets) and $T(A) = B$ is continuous, arc preserving and irreducible, T does not necessarily reduce to a homeomorphism. However, in this paper it will be shown, among other things, that if A is locally connected, if $T(A) = B$ is continuous and either A or B is cyclic (i. e., without cut points), then in order that T reduce to a homeomorphism it is necessary and sufficient that T be arc preserving and irreducible.

In this paper all transformations used will be supposed single valued and continuous.

It is readily seen that if A is a compact continuum and $T(A) = B$ is continuous then *there always exists a subcontinuum A_1 of A such that $T(A_1) = B$ and T is irreducible on A_1.* This follows at once from the Brouwer Reduction Theorem, since the property of being a subcontinuum of A mapping onto all of B under T clearly is inducible.

2. *Irreducible and arc-preserving transformations on locally connected continua.* Throughout this section we shall suppose A to be a compact locally connected continuum.

(2.1) *If A is a locally connected continuum and $T(A) = B$ is irreducible and arc preserving, then T is a homeomorphism on each true cyclic element of A on which T is not constant.*

Proof. Let E be such a true cyclic element of A. We have to prove that for each $b \in B$, $E \cdot T^{-1}(b)$ contains at most one point. Let us suppose this is not so. Then for some $b \in B$ there exists a non-degenerate arc $xpy \subset E$ such that $xpy \cdot T^{-1}(b) = x + y$. Now $T(xpy)$ is an arc ub and we may suppose p chosen so that $T(p) = u$. Let U_x and U_y be disjoint connected neighborhoods of x and y in A such that $T^{-1}(u) \cdot (U_x + U_y) = 0$ and $py \cdot U_x = xp \cdot U_y = 0$. Since there are non-cut points of A in both U_x and U_y, it follows from the irreducibility of T that there exist points $c \in U_x$ and $d \in U_y$ such that $T^{-1}[T(c)] \subset U_x$, $T^{-1}[T(d)] \subset U_y$. For if p is a non-cut point of A in U_x, say, there exists a neighborhood U_p of p such that $A - A \cdot U_x$ is contained in a single component N of $A - A \cdot U_p$; and since T is irreducible, there must exist a point c in U_x such that $N \cdot T^{-1}[T(c)] = 0$ and hence so that $T^{-1}[T(c)] \subset U_x$. Now let cq and dr be arcs in U_x and U_y respectively such that $cq \cdot xpy = q$, $dr \cdot xpy = r$. Let α denote the arc $cq + qpr + rd$.

We shall show that $T(\alpha)$ is not a simple arc. To do this it suffices to show that no one of the three points $T(c)$, $T(d)$, u separates the other two on $T(\alpha)$. Let β denote the subarc of the arc $T(qpr)$ from $T(q)$ to $T(r)$. Then since $T(q) \neq u \neq T(r)$, β does not contain u. Also $T(cq) \cdot u = T(dr) \cdot u = 0$. Thus $T(cq) + \beta + T(dr)$ is a connected subset of $T(\alpha)$ not containing u. But $T(cq + qp)$ is a connected subset of $T(\alpha)$ not containing $T(d)$ and $T(dr + rp)$ is a connected subset of $T(\alpha)$ not containing $T(c)$. Thus $T(\alpha)$ cannot be a simple arc, and our supposition that our theorem is not true leads to a contradiction.

(2.2) *Let $T(A) = B$ be irreducible and arc preserving, where A is locally connected. Then if B is cyclic, so also is A.*

Proof. Suppose, on the contrary, that A has a cut point p. Let us write $A = X + Y$ where X and Y are continua and $X \cdot Y = p$. Let $P = T^{-1}T(p)$ and $K = T^{-1}T(X)$.

(i) If $x \in Y \cdot (K - P)$, then for any arc px in A we have $px \subset K$.

For suppose not. Then for some $x' \in X$ we have $T(x') = T(x)$. Let $x'p$ be an arc in X. Then $T(x'p)$ is an arc in $T(X)$ and $T(px)$ is an arc in A not lying wholly in $T(X)$. Since each of the arcs $T(x'p)$ and $T(px)$ contains the two points $T(p)$ and $T(x)$, it follows that their sum contains a simple

closed curve. But $T(px) + T(x'p) = T(x'px)$ and this set must be a simple arc.

(ii) *If* $q \in (P - p)$, then for any arc pq in A we have either $pq \subset K$ or $pq \cdot K \subset P$.

For suppose pq contans a point x of $K - P$ and also a point y of $A - K$. Then by (i) we have $px \subset K$. Thus $y \subset xq$. Hence $T(px)$ is an arc in $T(X)$ and $T(xq)$ is an arc in B not lying wholly in $T(X)$. Since each of these arcs contains the two points $T(x)$ and $T(p)$ $[= T(q)]$, it follows that their sum contains a simple closed curve. This is impossible since $T(px) + T(xq) = T(pxq)$.

(iii) *If* R *is any component of* $A - K$, *then* $T[F(R)]$ *reduces to a single point.*[1]

For if not, there would exist an arc xy such that $T(x) \neq T(y)$, $x + y \subset K$ and $xy - (x + y) \subset R$. Now either $T(x)$ or $T(y)$ is $\neq T(p)$. Suppose $T(x) \neq T(p)$. Then by (i) any arc px in A is $\subset K$. Since then $px + xy$ is an arc pxy, it follows by (i) that $y \in P$. But since pxy contains $x \in (K - P)$ and also contains points in the open arc xy which are not in K, this contradicts (ii).

(iv) Now let R be a component of $A - K$ and let Q be the sum of all those components S of $A - K$ such that $T[F(S)] = T[F(R)] = a$. Then since $Q + T^{-1}(a)$ is closed, it follows that $T(Q) + a$ is closed. Since a is not a cut point of B it results that some point x of $T(Q)$ must be a limit point of $B - T(Q)$. Since Q is open in A, it follows from this that $T^{-1}(x)$ must intersect some component U of $A - K$ which does not belong to Q. Let $b = T[F(U)]$.

Now we have two essential cases to consider as follows: (I) $a \neq T(p) \neq b$, and (II) $a = T(p) \neq b$. In either case let $x_1 \in T^{-1}(x) \cdot Q$ and $x_2 \in T^{-1}(x) \cdot U$, let x_1a_1 and x_2b_1 be arcs such that $x_1a_1 - a_1 \subset Q$, $x_2b_1 - b_1 \subset U$, $a_1 \in T^{-1}(a)$, $b_1 \in T^{-1}(b)$.

In case I, we have $a_1, b_1 \in (K - P)$. Hence it follows by (i) that there exists an arc a_1b_1 in K. Then $T(a_1b_1)$ is an arc in $T(X)$ containing both a and b. But since $T(x_1) = T(x_2) = x$, it follows that $T(x_1a_1) + T(x_2b_1)$ contains an arc from a to b which clearly cannot lie in $T(X)$. Hence $T(a_1b_1) + T(x_1a_1) + T(x_2b_1) = T(x_1a_1 + a_1b_1 + b_1x_2) = T(x_1a_1b_1x_2)$ contains a simple closed curve, which is impossible.

In case II, consider an arc a_1p in A. If this arc contains b_1, then by (ii) we have $a_1p \subset K$ so that a_1p would contain a subarc a_1b_1 contained in K.

---

[1] $F(R)$ denotes the boundary of R relative to A, i.e., the set $\bar{R} - R$.

If a_1p does not contain b_1, then by (i) any arc pb_1 in A is $\subset K$; and $a_1p + pb_1$ contains an arc a_1b_1 some subarc a'_1b_1 of which is $\subset K$, where $a'_1 \epsilon P = T^{-1}(a)$. Let us agree to set $a_1 = a'_1$ in case $a_1b_1 \subset K$. Then $T(a'_1b_1)$ is an arc in $T(X)$ containing both $T(a'_1) = a$ and $T(b_1) = b$. Since $T(x_1) = T(x_2) = x$, it follows that $T(x_1a'_1) + T(x_2b_1)$ contains an arc from $T(a'_1)$ to $T(b_1)$ which does not lie wholly in $T(X)$. Hence

$$T(x_1a'_1) + T(a'_1b_1) + T(b_1x_2) = T(x_1a'_1 + a'_1b_1 + b_1x_2) = T(x_1a'_1b_1x_2)$$

contains a simple closed curve, which is impossible.

(2.21) Corollary. *Let $T(A) = B$ be arc preserving, where A is locally connected and B is cyclic. If A has a cut point p, then for each component C of $A - p$ we have $T(C + p) = B$.*

This results at once from the proof of (2.2), because we can take X in that proof to be $C + p$; and it will be noted that the irreducibility of T was used only once, namely, in (iv) to insure the existence of at least one component R of $A - T^{-1}T(X)$.

(2.3) *If A is a locally connected continuum and $T(A) = B$ is irreducible and arc preserving and if for some $b \epsilon B$, x, $y \epsilon T^{-1}(b)$, $x \neq y$, then on any true cyclic element in the cyclic chain[2] $C(x, y)$, T must be constant.*

Proof. Let xpy be any arc in A from x to y. Now if $T(xpy) = b \epsilon B$, our result follows from (2.1). Thus we suppose $T(xpy)$ contains more than one point. Also, without loss of generality, we can assume that $xpy \cdot T^{-1}(b) = x + y$. Hence $T(xpy)$ is an arc ub; and clearly we may suppose p chosen so that $T(p) = u$.

Now if we suppose (2.3) does not hold, there exists a true cyclic element E of A intersecting xpy in a non-degenerate arc cd where we have the order x, c, d, y and where T is not constant on E. Since by (2.1) T must be a homeomorphism on cd and since $T(xp) = T(py) = ub$, it follows that either $cd \subset xp$ or $cd \subset py$. Let us suppose the latter. Then we have the order x, p, c, d, y on xpy, where some points may coincide. Now there must exist a point w on $py - (p + y)$ which is a limit point of $E - cd$ and such that $u \neq T(w) \neq b$. For if not, then E would contain an arc $p'zy'$ such that $p'zy' \cdot xpy = p' + y'$ and $T(p') = u$, $T(y') = b$; but this is impossible,

[2] If M is a locally connected continuum, a subset of M is called an A-set provided it is closed and it contains every simple arc in M whose endpoints lie in it. If $x, y \epsilon M$, the smallest A-set in M containing $x + y$ is called the *cyclic chain* $C(x, y)$. See my paper in *American Journal of Mathematics*, vol. 50 (1928), pp. 167-194; see also Kuratowski and Whyburn, *Fundamenta Mathematicae*, vol. 16 (1930), pp. 305-331.

because then $T(xp' + p'zy') = T(p'y') + T(p'zy')$, which is a simple closed curve since $p'y' + p'zy'$ is a simple closed curve and T is a homeomorphism on every such curve by (2.1). (Note: $p'y'$ denotes the arc $p'y'$ of xy). Thus such a point w exists.

It follows readily [see the proof of (2.1)] that there exists an arc rv in E such that

$$rv \cdot xpy = v, \quad T(rv) \cdot (u + b) = 0, \quad \text{and} \quad T(r) \cdot ub = 0.$$

Now $xpv + rv$ is an arc in A; but $T(xpv) = T(xp) = ub$, whereas $T(rv)$ is an arc in B intersecting ub but containing neither u nor b and yet not being contained in ub. Thus $T(xpv + rv)$ cannot be a simple arc, contrary to hypothesis.

(2.4) *If A is locally connected and $T(A) = B$ is irreducible and arc preserving, then for each true cyclic element E_b in B there exists a unique true cyclic element E_a in A such that $T(E_a) = E_b$.*

Proof. Let $K = T^{-1}(E_b)$.

(i) For each component R of $A - K$, we have $T[F(R)] = p \epsilon E_b$.

For suppose, on the contrary, that for some $x, y \epsilon F(R)$ we have $T(x) \neq T(y)$. Then $T(R)$ is a connected subset of $B - E_b$ having at least the two limit points $T(x)$ and $T(y)$ in E_b. Clearly this is impossible since E_b is a cyclic element of B.

(ii) If we define the transformation $Z(A) = E_b$ as follows:

$$Z(x) = T(x) \text{ for } x \epsilon K, \qquad Z(x) = T[F(R)] \text{ for } x \epsilon R,$$

where R is a component of $A - K$, then Z is continuous and arc preserving.

The continuity is immediate. Let pq be any arc in A. Then $T(pq)$ is an arc ab in B. If $ab \cdot E_b = 0$, then $pq \subset$ some R; whence $Z(pq) = T[F(R)]$ = a single point. If $ab \cdot E_b \neq 0$, then $ab \cdot E_b$ is a subarc $a'b'$ of ab which may reduce to a single point. Let $x \epsilon (pq - K \cdot pq)$. Then $Z(x) = T[F(R)]$, where R is the component of $A - K$ containing x, and clearly $Z(x) \epsilon a'b'$. Thus $Z(pq) = a'b'$ and Z is arc preserving.

Let E_a be an A-set in A which intersects every one of the sets $T^{-1}(b)$, $b \epsilon E_b$, and is irreducible with respect to the property of being an A-set intersecting every one of these sets. Then $Z(E_a) = E_b$ and Z is arc preserving on E_a since it is arc preserving on A. It now follows from (2.21) that E_a can have no cut point, and hence E_a is a true cyclic element of A. But T is a homeomorphism on each true cyclic element of A on which it is not constant, by (2.1), and hence is a homeomorphism on E_a; and since $T(E_a) \supset E_b$, this gives $T(E_a) = E_b$ as was to be proven.

Finally, there can exist no other true cyclic element $E_{a'}$ of A such that $T(E_{a'}) = E_b$. For if so, let x be a point of E_b such that $T^{-1}(x) \cdot [E_a + E_{a'}]$ contains no cut point of A. Then if $x_1 = E_a \cdot T^{-1}(x)$, $x_2 = E_{a'} \cdot T^{-1}(x)$, we have $C(x_1, x_2) \supset E_a + E_{a'}$ and this contradicts (2.3).

3. *Recapitulation. Conditions for homeomorphism.* The results just established in § 2 yield at once the following theorems.

(3.1) THEOREM. *If A is a compact locally connected continuum and $T(A) = B$ is irreducible and arc preserving, then T maps the true cyclic elements E_a of A on which T is not constant into the true cyclic elements E_b of B in a $(1-1)$ way and T is a homeomorphism on each E_a. (i. e., for each such E_a, $T(E_a)$ is an E_b and for each E_b there is a unique E_a such that $T(E_a) = E_b$).*

(3.2) THEOREM. *Let A be a compact locally connected continuum, let $T(A) = B$ be continuous and suppose that either A or B is cyclic. Then in order that T be a homeomorphism it is necessary and sufficient that it be irreducible and arc preserving.*

It is clear that for the validity of the former of these theorems the irreducibility of T is indispensable. However, if in the latter theorem we specify that B is cyclic, it seems possible that the condition of irreducibility on T could be dispensed with. This is indeed the case, as will now be shown, provided we assume further that A is hereditarily locally connected, i. e., that not only A but also every subcontinuum of A is locally connected.

(3.3) THEOREM. *Let A be hereditarily locally connected, let $T(A) = B$ be continuous and suppose that B is cyclic. Then in order that T be a homeomorphism it is necessary and sufficient that it be arc preserving.*

Proof. The necessity of the condition is obvious. We proceed to establish the sufficiency. In view of (3.2) it suffices to prove that T is irreducible. Let us suppose, on the contrary, that for some proper subcontinuum A' of A we have $T(A') = B$. We may suppose A' chosen so that T is irreducible on A'. Also, since T is arc preserving on A, clearly it is likewise arc preserving on A'. Furthermore, since A' is locally connected and B is cyclic, it follows by (3.2) that A' is cyclic and T is a homeomorphism on A'.

Now since $A' \neq A$, there exists an arc ab in A such that $ab \cdot A' = a + b$. Then $T(ab)$ is an arc and, since T is a homeomorphism on A', so also is $A' \cdot T^{-1}T(ab)$. Call this latter arc xy. Since A' is cyclic, there exists in A' an arc $\alpha\beta$ such that $xy \cdot \alpha\beta = \alpha + \beta$ and we have the order $x \leq \alpha < \beta \leq y$ on xy. Let $\alpha z \beta$ be the subarc of xy from α to β.

(i) If uv is any proper subarc of the arc $T(ab) = rs$, there exists a proper subarc pq of ab such that $T(pq) \supset uv$.

For let m be a point of $rs - uv$ such that $T(a) \neq m \neq T(b)$. Then either rm or ms contains uv, suppose the former. Then since $T^{-1}(m) \cdot (a + b) = 0$, ab contains a proper subarc pq where $p \in T^{-1}(m)$, $q \in T^{-1}(r)$; and since $T(pq)$ is a subarc of rs containing both r and m, we have $T(pq) \supset uv$. Similar reasoning applies in case $ms \supset uv$.

(ii) If $\alpha z\beta \neq xy$, then $T(\alpha z\beta)$ is a proper subarc uv of rs (since T is a homeomorphism on A'). Thus by (i) we have a proper subarc pq of ab such that $T(pq) \supset uv$. Not both a and b can belong to pq. Suppose b non-$\in pq$ and suppose we have the order a, p, q, b on ab. Now $b \in xy$ and xy contains either an arc $b\beta$ not containing α or an arc $b\alpha$ not containing β. Call this arc t. Then in either case it is apparent that $\alpha\beta + t + pqb$ is a simple arc γ. But $T(\gamma) \supset T(\alpha\beta) + T(pq) = T(\alpha\beta) + uv = T(\alpha\beta) + T(\alpha z\beta)$; and this latter set is a simple closed curve, since $\alpha\beta + \alpha z\beta$ is a simple closed curve in A' and T is a homeomorphism on A'. Thus this case leads to a contradiction.

(iii) If $\alpha z\beta = xy$ we may suppose we have $\alpha = x$, $\beta = y$. Now both a and b lie on xzy and we may suppose we have the order $x \leq a < z < b \leq y$, on xzy. Since $T(a) \neq T(b)$, there exist arcs ac and db in ab and zb respectively such that $T(ac) \cdot T(db) = 0$. Since $T(ab) = T(xzy) = rs$, this gives $T(by) + T(xazd) \supset T(ac)$. Whence $T(by) + T(xazd) + T(cb) = rs$. Thus $T(cb + by + \beta\alpha + xazd) = T(cb) + T(xazd) + T(by) + T(\alpha\beta) = rs + T(\alpha\beta)$. But this is impossible, since $cb + by + \beta\alpha + xazd$ is a simple arc whereas $rs + T(\alpha\beta)$ is a simple close curve. This contradiction completes our proof.

4. *A-set reversing transformations.* If A is a compact locally connected continuum and $T(A) = B$ is continuous (as we shall suppose throughout this section), then T is said to be *A-set reversing* provided that for each $b \in B$, $T^{-1}(b)$ is either a single point or an A-set in A.

(4.1) THEOREM. *In order that T be A-set reversing it is necessary and sufficient that T be monotone*[3] *on every simple arc in A.*

To prove the condition necessary, suppose T *A-set* reversing, let pq be any simple arc in A and let $b \in T(pq)$. Then since $T^{-1}(b)$ is either a single point or an A-set, it follows that $pq \cdot T^{-1}(b)$ is[4] in either case a connected set. Consequently T is monotone on pq (indeed on any connected subset of A).

[3] A transformation $T(X) = Y$ is *monotone* provided that for each $y \in Y$, $T^{-1}(y)$ is connected. See C. B. Morrey, *American Journal of Mathematics*, vol. 57 (1935), pp. 17-50.

[4] For properties of A-sets, see Kuratowski and Whyburn, *loc. cit.*

The condition is also sufficient. For suppose it satisfied and let $b \in B$. Let $p, q \in T^{-1}(b)$ and let pq be any arc in A from p to q. Since by hypothesis T is monotone on pq and since $T(p) = T(q) = b$, we must have $T(pq) = b$ or $pq \subset T^{-1}(b)$. Thus $T^{-1}(b)$ is either a single point or an A-set and T is A-set reversing.

(4.11) Corollary. *Any A-set reversing transformation on A is arc preserving.*

For the property of being a simple arc is invariant under monotone transformations.

(4.12) Corollary. *If T is A-set reversing on A, then T is monotone on every connected subset of A.*

(4.13) *Any monotone transformation on a dendrite is arc preserving.*

(4.14) *Any monotone transformation on A is arc preserving on the set $K + H$ of all cut points and all end points of A.*

(4.2) *If A is a locally connected continuum and $T(A) = B$ is A-set reversing, then for every A-set A_b in B, $T^{-1}(A_b)$ is an A-set in A (or a single point).*

Proof. Let $p, q \in T^{-1}(A_b)$, $p \neq q$, and let pq be any arc in A from p to q. If $T(p) = T(q)$, then $T(pq) = T(p) \in A_b$ so that $pq \subset T^{-1}(A_b)$, since T is A-set reversing. If $T(p) \neq T(q)$, then by (4.11), $T(pq)$ is an arc, from $T(p)$ to $T(q)$. Whence $T(pq) \subset A_b$ so that $pq \subset T^{-1}(A_b)$. Thus $T^{-1}(A_b)$ is an A-set.

(4.3) *If A is a locally connected continuum and $T(A) = B$ is A-set reversing, then for each true cyclic element E_b in B there exists a unique true cyclic element E_a in A such that $T(E_a) = E_b$ and T is a homeomorphism on E_a.*

Proof. Since T is monotone, there exists [5] a true cyclic element E_a of A such that $T(E_a) \supset E_b$. Now by (4.2), $T^{-1}(E_b)$ is an A-set in A. Since $T^{-1}(E_b)$ contains at least two points of E_a, this gives $E_a \subset T^{-1}(E_b)$. Whence $T(E_a) \subset E_b$. Accordingly $T(E_a) = E_b$. Finally, for no $b \in B$ can $T^{-1}(b)$ contain more than one point of E_a; for if it did, we would have $E_a \subset T^{-1}(b)$, which is impossible. Thus T is a homeomorphism on E_a. Obviously E_a is uniquely determined.

(4.31) Corollary. *If A and B are locally connected and cyclic and $T(A) = B$ is A-set reversing, T is a homeomorphism.*

The University of Virginia.

[5] See my paper in the *American Journal of Mathematics*, vol. 56 (1934), pp. 133-146.

AN ANALYTIC CHARACTERIZATION OF SURFACES OF FINITE LEBESGUE AREA. PART II.

By Charles B. Morrey, Jr.[1]

In Part I of this paper,[2] the author gives an analytic characterization of surfaces of finite Lebesgue area in which only "non-degenerate" surfaces are admitted for discussion. The object of the present part is to generalize these results to arbitrary surfaces of finite area, the hypothesis of non-degeneracy being dropped.

Throughout this paper, we shall use the following vector notation: the letters x and X shall stand for the vector of the coördinates $(x^1, \cdots, x^N)$ and $(X^1, \cdots, X^N)$ of a point in the x-space in which a given surface will lie, the letters u and U for the coördinates of a point of the space containing the set on which the given surface will be parametrically represented, the sum and difference of pairs of these letters will denote the sum and difference vectors, x_α and $\partial x/\partial \alpha$ for the vector $(\partial x^1/\partial \alpha, \cdots, \partial x^N/\partial \alpha)$, α being a parameter, $x(u)$ and $X(U)$ will be vector functions, and if ϕ is a vector in any space, $|\phi|$ will denote its length. We may write $x(P)$ to mean $x(u)$ where u is the coördinate vector of P, u_P for the coördinate vector of P, etc. Given a point set E, $\bar{E}$ will denote its closure and E^* the set of its frontier points. All vector functions occurring in a transformation or the representation of a surface will be assumed continuous. As there are a great many topological terms used whose definitions are long, we shall merely give the reference to where they may be found in the author's paper, "The topology of path surfaces,"[3] which we shall hereafter refer to by the letter T. as frequent references to this paper will be made.

3. *The existence of a "generalized conformal" map of an arbitrary surface of finite Lebesgue area on a hemicactoid* (T. § 2, def. 12).

Definition 1. Let $\bar{H}$ be a hemicactoid and $\bar{h}$ a continuous curve (T. § 1, Definition 9) which is a subset of $\bar{H}$. Let $\bar{b}$ consist of all the points of $\bar{h}$ not separated from $\bar{B}$ in $\bar{H}$ (T. § 1, Definition 4) by any point of $\bar{h}$; the set $\bar{b}$ is a continuous curve which is either a subset of $\bar{B}$ or a subset of a single cyclic

[1] National Research Fellow (1931-33).

[2] *American Journal of Mathematics*, vol. 57 (1935), pp. 692-702.

[3] *American Journal of Mathematics*, vol. 57 (1935), pp. 17-50.

element (T. § 1, Definition 13) of $\bar{H}$ which does not contain all of that element unless that element is a point, in which case $\bar{b}$ is that point. If $\bar{b}$ is a subset of $\bar{B}$ which is a base set (not necessarily canonical) (T. § 2, Definitions 6 and 11), is a single point, or is a subset of a double cone (T. § 2, Definition 7) of $\bar{H}$ and does not separate that double cone being therefore homeomorphic with a base set, and if $\overline{(\bar{h} - \bar{b})}$ is a sum of canonical cactoids (T. § 2, Definitions 10 and 11), then we say that $\bar{h}$ is a *sub-hemicactoid of* $\bar{H}$ *and* $\bar{b}$ *is its base set.* $\bar{h}$ is not necessarily a hemicactoid in the ordinary sense as $\bar{b}$ may not be a canonical base set although it is always homeomorphic to one.

Definition 2. Let S be a surface represented by $x = X(U)$ on a hemicactoid $\bar{H}$ and let $\bar{h}$ be a sub-hemicactoid of $\bar{H}$ with base set $\bar{b}$. From T. § 4, Theorem 1 and the above remarks it follows that there exists a monotone transformation (T. § 2, Definition 5) $U = U(u)$, with corresponding collection of continua G (see T. § 2, Theorem 4), of a Jordan region $\bar{r}$ into $\bar{h}$ which carries the subcollection G_0 of T. § 2, Lemma 8, topologically (T. § 2, Definition 4 and Theorem 2) into $\bar{b}$. If we define $x(u) = X[U(u)]$, $u \varepsilon \bar{r}$, then $x = x(u)$ is a surface and all surfaces thus obtained are identical. We call this surface the *sub-surface of* S *corresponding to* $\bar{h}$.

THEOREM 1. *Let* S *be a surface represented on a hemicactoid* $\bar{H}$ *by* $x = X(U)$. *Let* $\bar{h}_1$ *and* $\bar{h}_2$ *be two sub-hemicactoids of* $\bar{H}$ *such that* $\bar{h}_1 + \bar{h}_2 = \bar{H}$ *and* $\bar{h}_1 \cdot \bar{h}_2$ *is a single cut point of* $\bar{H}$. *Then if* S_1 *and* S_2 *are the sub-surfaces of* S *corresponding, respectively, to* $\bar{h}_1$ *and* $\bar{h}_2$, *we have*

$$L(S) = L(S_1) + L(S_2).^{4}$$

Proof. Suppose $\bar{h}_1 \cdot \bar{h}_2$ is not the whole base set of either $\bar{h}_1$ or $\bar{h}_2$. Then $\bar{b} \cdot \bar{b}_1$ and $\bar{b} \cdot \bar{b}_2$ each contain more than one point and $\bar{h}_1 \cdot \bar{h}_2 = \bar{b}_1 \cdot \bar{b}_2 \varepsilon \bar{b}$. Now, we may find a monotone transformation with corresponding collection G_1 of $\bar{r}_1$, $\bar{r}_1 : 0 \leqq u \leqq 3/8$, $0 \leqq v \leqq 1$, into $\bar{h}_1$ which carries the subcollection $G_{1,0}$ into $\bar{b}_1$, the segment, $u = 3/8$, $0 \leqq v \leqq 1$, being carried into the point $\bar{h}_1 \cdot \bar{h}_2$. Similarly we may carry $\bar{r}_2$, $\bar{r}_2 : 5/8 \leqq u \leqq 1$, $0 \leqq v \leqq 1$, on $\bar{h}_2$, the segment, $u = 5/8$, $0 \leqq v \leqq 1$, being also carried into $\bar{h}_1 \cdot \bar{h}_2$. In this case we unite these to give a map $U = U(u)$ of $\bar{H}$ on $\bar{Q}$ by first defining $U(u) = U_{\bar{h}_1 \cdot \bar{h}_2}$, $u \varepsilon \gamma$, $\gamma : 3/8 \leqq u \leqq 5/8$, $0 \leqq v \leqq 1$, and then defining $U(u)$ in $\bar{r}_1$ and $\bar{r}_2$ by the condition that $U = U(u)$ give the above determined mappings. Clearly $U = U(u)$ carries $G = G_1 + G_2$ topologically into $\bar{H}$ and its subcollection G_0 topologically into $\bar{B}$.

---

[4] $L(S)$ denotes the Lebesgue area (see introduction to part I) of S.

If $\bar{h}_1 \cdot \bar{h}_2$ is the whole base set of one of the $\bar{h}_i$, we shall call that one $\bar{h}_2$. We may map $\bar{h}_1$ on $\bar{Q}$ by $U = \tilde{U}(u)$ as $\bar{h}_1$ was mapped on $\bar{r}_1$ above except that we require that a single point u_1 of $\bar{Q}$ be carried into $\bar{h}_1 \cdot \bar{h}_2$. We extend $\tilde{U}(u)$ to the square $\bar{Q}_\lambda$, $\bar{Q}_\lambda\colon -\lambda \leqq u,\ v \leqq 1 + \lambda$ $(\lambda > 0)$, by defining $\tilde{U}(u) = \tilde{U}(u_0)$ for $u \in \bar{Q}_\lambda - \bar{Q}$, u_0 being the point of intersection of the ray, $(1/2, 1/2)$, (u, v) with Q^*. Now u_1 is certainly interior to $\bar{Q}$. Let $u' = U'(u)$ be the radial contraction of $\bar{Q}_\lambda$ into $\bar{Q}$ with center at $(1/2, 1/2)$ and define $u'_1 = u'(u_1)$. Define the transformation (monotone but discontinuous at u'_1) $u'' = u''(u')$ of $\bar{Q}$ into $\bar{r}_1 + \gamma$, $\bar{r}_1$ being the doubly connected region consisting of all points $u'_1 + \lambda(\bar{u}'_0 - u'_1)$, $1/2 \leqq \lambda \leqq 1$, $\bar{u}'_0$ denoting any point of Q^*, by

$$u''(u') = u'_0 + (u' - u'_0)/2, \qquad u' \neq u'_1,$$

u'_0 being the last point of $\bar{Q}$ on the ray $u'_1 u'$; u'_1 is carried into all the points of the region γ, consisting of all points $u'_1 + \lambda(u'_0 - u'_1)$, $3/8 \leqq \lambda \leqq 1/2$, u'_0 ranging over all of Q^*. The transformation from u'' to u is continuous: $u = u(u'')$. Now let $r_2 = Q - (\bar{r}_1 + \gamma)$; we can map $\bar{h}_2$ on $\bar{r}_2$ by a monotone transformation $U = U(u'')$ which carries r^*_2 into $\bar{h}_1 \cdot \bar{h}_2$. If we define $U(u'') = \tilde{U}[u(u'')]$ for $u'' \in \bar{r}_1 + \gamma$, then $U = U(u'')$, in this case also, gives a monotone transformation, with corresponding collection G, of $\bar{Q}$ into $\bar{H}$ which carries G_0 into $\bar{B}$.

Now in either case above S is represented on $\bar{Q}$ by $x = x(u) = X[U(u)]$. Let $x_0 = x(u_0)$, $u_0 \in \gamma$. Clearly, in either case above, the surfaces S_i are represented on $\bar{Q}$ by $x = x_i(u)$ $(i = 1, 2)$, where

$$\begin{aligned} &x_1(u) = x(u), u \in \bar{r}_1; \ x_1(u) = x_0, u \in \gamma; \ x_1(u) = x_0, \quad u \in \bar{r}_2; \\ &x_2(u) = x_0, \quad u \in \bar{r}_1; \ x_2(u) = x_0, u \in \gamma; \ x_2(u) = x(u), u \in \bar{r}_2. \end{aligned}$$

Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n\colon x = \bar{x}_n(u)$, $u \in \bar{Q}$, be a sequence of polyhedra approaching S, whose areas approach $L(S)$, the $\bar{x}_n(u)$ being linear in patches in $\bar{Q}$ and approaching $x(u)$ uniformly. Let $\{r_m\}$ and $\{R_m\}$ be two sequences of numbers approaching zero so that r_m/R_m approaches zero. For each m, let n_m be $\geqq m$ and so large that $|\,\bar{x}_{n_m}(u) - x_0\,| < r_m/2$, $u \in \gamma$. Now define: $\hat{x}_m(u) = \bar{x}_{n_m}(u)$ for those values of u in $\bar{Q}$ for which $|\,\bar{x}_{n_m}(u) - x_0\,| \geqq R_m$, $x_m(u) = x_0$ for those values of u in Q for which $|\,\bar{x}_{n_m}(u) - x_0\,| \leqq r_m$ (these constituting a finite number of regions bounded by a finite number of curves analytic in pieces, one of these regions containing γ), and

$$\hat{x}_m(u) = x_0 + [(|\,\bar{x}_{n_m}(u) - x_0\,| - r_m)/(R_m - r_m)][\bar{x}_{n_m}(u) - x_0]$$

for the remaining values of u. This function $\hat{x}_m(u)$ is analytic in patches and equal to x_0 in the set in parenthesis above containing γ. Furthermore it is immediate that $L(\hat{S}_m) \leqq (R_m/(R_m - r_m)) \cdot L(\bar{\Pi}_{n_m})$. But now it is

well known that, in such a surface $\hat{S}_m$, we may inscribe a polyhedron $\bar{\Pi}$, $\bar{\Pi}: x = x(u)$ so that $\| S_m, \bar{\Pi} \|$ and $| L(S_m) - L(\bar{\Pi}) |$ are arbitrarily small and $\bar{x}(u) = x_0$, $u \varepsilon \gamma$. Hence it is clear that we may replace the sequence $\{\bar{\Pi}_n\}$ by a sequence $\{\Pi_m\}$, $\Pi_m : x = x_m(u)$, $u \varepsilon \bar{Q}$, where

$$\lim_{m\to\infty} \| \Pi_m, S \| = \lim_{m\to\infty} | L(\Pi_m) - L(S) | = 0,$$

the $x_m(u)$ being linear in patches, converging uniformly to $x(u)$, and with $x_m(u) = x_0$, $u \varepsilon \gamma$. Then if we define, for each m, $x_{1,m}(u)$ and $x_{2,m}(u)$ from $x_m(u)$ as $x_1(u)$ and $x_2(u)$ were formed from $x(u)$, it is clear that $L(\Pi_m) = L(\Pi_{1,m}) + L(\Pi_{2,m})$ and that $\{\Pi_{i,m}\}$ converges to S_i, $\Pi_{i,m} : x = x_{i,m}(u)$ $(i = 1, 2)$. Thus $L(S) \geqq L(S_1) + L(S_2)$.

Now, let $\{\bar{\Pi}_{i,n}\}$, $x = \bar{x}_{i,n}(u)$, $u \varepsilon \bar{Q}$, be sequences approaching S_i so that $L(\bar{\Pi}_{i,n})$ approaches $L(S_i)$, each $\bar{x}_{i,n}(u)$ being linear in patches over $\bar{Q}$ and the $\bar{x}_{i,n}(u)$ converging uniformly to $x_i(u)$ $(i = 1, 2)$. Repeat the above process on each sequence obtaining sequences $\{\Pi_{i,m}\}$, $\Pi_{i,m} : x = x_{i,m}(u)$, where $\Pi_{i,m}$ converges to S_i, $L(\Pi_{i,m})$ to $L(S_i)$, $x_{i,m}(u)$ uniformly to $x_i(u)$, each $x_{i,m}(u)$ being linear in patches on $\bar{Q}$ and equal to x_0 on $\bar{r}_{3-i} + \gamma$ $(i = 1, 2)$. Then define (for each m) Π_m by $\Pi_m : x = x_m(u)$ where

$$x_m(u) = x_{1,m}(u), u \varepsilon \bar{r}_1; \quad x_m(u) = x_0, u \varepsilon \gamma; \quad x_m(u) = x_{2,m}(u), u \varepsilon \bar{r}_2.$$

Then clearly $L(\Pi_m) = L(\Pi_{1,m}) + L(\Pi_{2,m})$ and $\lim_{m\to\infty} \| \Pi_m, S \| = 0$. Thus $L(S) \leqq L(S_1) + L(S_2)$ which completely demonstrates the theorem.

THEOREM 2. *Let S be a surface represented by $x = x(u)$ on a Jordan region $\bar{r}$, $x(u)$ being constant over r^*. Then there exists a sequence $\{\Pi_m\}$, $\Pi_m : x = x_m(u)$, of polyhedra* (1) *where $x_m(u)$ is non-degenerate except that it is constant over r^*,* (2) $\lim_{m\to\infty} \| S, \Pi_m \| = 0$, *and* (3) $L(S) = \lim_{m\to\infty} L(\Pi_m)$.

Proof. Take $\bar{r} = \bar{Q}$. Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n : x = \bar{x}_n(u)$, $u \varepsilon \bar{Q}$, be any sequence of polyhedra approaching S such that $L(\bar{\Pi}_n)$ approaches $L(S)$, the $\bar{x}_n(u)$ being linear in patches on $\bar{Q}$ and $\bar{x}_n(u)$ converging uniformly to $x(u)$. Let $x_0 = x(u_0)$, $u_0 \varepsilon Q^*$. By the process of Theorem 1, using this x_0, we may replace $\{\bar{\Pi}_n\}$ by a sequence $\{\bar{\Pi}_m\}$, $\bar{\Pi}_m : x = \bar{x}_m(u)$, satisfying (2) and (3), the $\bar{x}_m(u)$ converging uniformly to $x(u)$, being linear in patches on $\bar{Q}$ and equal to x_0 on Q^*. By moving the vertices of each $\bar{\Pi}_m$ slightly, we can obtain a sequence $\{\Pi_m\}$ of the desired type.

LEMMA 1. *If S is a surface represented by $x = X(U)$ on a hemicactoid $\bar{H}$ consisting of a segment, then $L(S) = 0$ whether $\bar{B}$ is the whole segment or just one end point.*

Proof. *Case* I: $\bar{H} = \bar{B} =$ a segment, $0 \leqq u \leqq 1$, $v = 0$. Then if we define $x^i(u, v) = X^i(u)$, $x^i = x^i(u, v)$, $0 \leqq v \leqq 1$ is a representation of S on Q. It is clear that we may approximate to S by polyhedra of the same form which obviously have zero area. Thus $L(S) = 0$ in this case.

Case II: $\bar{H}$ a vertical segment $\xi = \eta = 0$, $0 \leqq \zeta \leqq 1$, $(0, 0, 0)$ being $\bar{B}$. Then S can be represented on $\bar{Q}$ by $x = x(u) = X[U(u)]$, $U(u)$ being defined by $\xi = \eta = 0$,

$$\zeta = 1 - \sqrt{[(u - 1/2)^2 + (v - 1/2)^2]/[(U - 1/2)^2 + (V - 1/2)^2]},$$

where (U, V) is the last point of $\bar{Q}$ on the ray, $(1/2, 1/2)$, (u, v). It is clear in this case also that S can be approximated to by a sequence of polyhedra each of area zero.

LEMMA 2. *If S is a surface represented on a simple cyclic chain $\bar{C}$ (of type A or B* [see T. § 2, Definition 9]), *then $L(S)$ is the sum of the areas of the subsurfaces of S corresponding to the non-degenerate cyclic elements of the chain.*

Proof. Suppose S is represented on $\bar{C}$ by $x = X(U)$. For each n, we know that the number of non-degenerate cyclic elements of $\bar{C}$ which are of diameter $\geqq 1/n$ is finite. Hence, for each n, form $\bar{C}_n$ from $\bar{C}$ by replacing each non-degenerate cyclic element of $\bar{C}$ of length $< 1/n$ by its axis. If $\bar{C}$ is of type B, $\bar{C}_n \subset \bar{C}$ and we define $X_n(U) = X(U)$ on $\bar{C}_n$. If $\bar{C}$ is of type A, define $X_n(U) = X(U)$ on $\bar{C} \cdot \bar{C}_n$ and $X_n(U) = X(U')$ for U on the axis of one of the replaced double cones of $\bar{C}$, U' ranging linearly over the broken line consisting of the two equal sides of a generating triangle of that double cone as U ranges over the axis. Define S_n by $S_n : x = X_n(U)$, $U \in \bar{C}_n$. By the proof of T. § 4, Lemma 5, we see that $\lim_{n\to\infty} \| S, S_n \| = 0$. Let $P_0, \cdots, P_{k_n}$ be the end points arranged in order, of $\bar{C}$ and of the cyclic elements of $\bar{C}$ which are of diameter $\geqq 1/n$ and let $\bar{C}^{(1)}, \cdots, \bar{C}^{(k_n)}$ and $\bar{C}_n^{(1)}, \cdots, \bar{C}_n^{(k_n)}$ be the parts into which $\bar{C}$ and $\bar{C}_n$, respectively, are divided by these points, the $S^{(i)}$ and $S_n^{(i)}$ being the corresponding subsurfaces of S and S_n. Each $\bar{C}_n^{(i)}$ is either a segment or a cyclic element of $\bar{C}$ of diameter $\geqq 1/n$. By repeated application of Theorem 1,

$$L(S) = \sum_{i=1}^{k_n} L(S^{(i)}), \qquad L(S_n) = \sum_{i=1}^{k_n} L(S_n^{(i)}).$$

Since $S_n^{(i)}$ is either identical with $S^{(i)}$ or corresponds to a segment of $\bar{C}_n$, in

which case $L(S_n^{(i)}) = 0$, it follows immediately that $L(S_n) \leqq L(S)$ and hence that $L(S)$ is equal to $\sum_{i=1}^{\infty} L(S'^{(i)})$ where the $S'^{(i)}$ are the subsurfaces of S corresponding to non-degenerate cyclic elements of $\bar{C}$.

LEMMA 3. *Let S be a surface represented by $x = x(u)$ on a Jordan region $\bar{r}$. Let γ be a maximal continuum over which $x(u)$ is constant which separates $\bar{r}$, and let d be a component of $\bar{r} - \gamma$ which may be a bounded complementary domain* (T. footnote p. 31) *of γ or a set consisting of a region, bounded by a portion of γ and a portion (possibly all) of r^*, plus that portion of r^* not in γ. Define $x_1(u) = x(u)$ for $u \in \bar{r} - d$, and $x_1(u) = x(u_0)$ for $u \in d$, u_0 being in γ. Then, if we define S_1 by $S_1 : x = x_1(u)$, $u \in \bar{r}$,*

$$L(S_1) \leqq L(S).$$

Proof. Take $\bar{r} = \bar{Q}$. Let $\{\bar{\Pi}_n\}$, $\bar{\Pi}_n : x = \bar{x}_n(u)$, $u \in \bar{Q}$, be a sequence of polyhedra approaching S such that $\lim_{n\to\infty} L(\bar{\Pi}_n) = L(S)$, the $\bar{x}_n(u)$ being linear in patches and converging uniformly to $x(u)$ on $\bar{Q}$. By the method of Theorem 1, we may replace $\{\bar{\Pi}_n\}$ by a sequence $\{\Pi_n\}$ of polyhedra, $\Pi_n : x = x_n(u)$, where $\lim_{n\to\infty} \| S, \Pi_n \| = \lim_{n\to\infty} | L(S) - L(\Pi_n) | = 0$, and where the $x_n(u)$ are linear in patches, approach $x(u)$ uniformly, and $x_n(u) = x_0 = x(u_0)$ in some closed connected region γ_n of finite connectivity, bounded by polygons, and including γ in its interior, the u_0 above being in γ.

Now, for each n, let d_n consist of all of the (finite number) of components of $\bar{Q} - \gamma_n$ which lie in d. Now, clearly, since γ is maximal, $\gamma = \prod_{n=1}^{\infty} \gamma_n$. Hence if, for each n, we define $x_{1,n}(u) = x_n(u)$, $u \in \bar{Q} - d_n$ and $x_{1,n}(u) = x_0$, $u \in d_n$, then the surfaces $\Pi_{1,n}$, $\Pi_{1,n} : x = x_{1,n}(u)$, are polyhedra approaching S_1 where also $L(\Pi_{1,n}) \leqq L(\Pi_n)$. Hence

$$L(S_1) \leqq \varliminf_{n\to\infty} L(\Pi_{1,n}) \leqq \lim_{n\to\infty} L(\Pi_n) = L(S)$$

which proves the lemma.

THEOREM 3. *If S is a surface represented on a hemicactoid $\bar{H}$, then $L(S)$ is the sum of the Lebesgue areas of the subsurfaces of S corresponding to the non-degenerate cyclic elements of $\bar{H}$.*

Proof. Let $\bar{H} = \sum_{m=1}^{\infty} \bar{B}_m + \sum_{n=1}^{\infty} \bar{C}_n + H^*$, where H^* is a completely disconnected set of end points of $\bar{H}$ which are limit points of $\Sigma \bar{B}_m + \Sigma \bar{C}_n$; the

$\bar{B}_m$ are the simple chains of type B of the base set and $\bar{C}_n$ are those of type A. We may arrange all of these into a single sequence $\{\bar{D}_n\}$ such that $\bar{D}_n \cdot (\bar{D}_1 + \cdots + \bar{D}_{n-1})$ is a single point. Let S be represented on $\bar{H}$ by $x = X(U)$ and let $U = U(u)$ be a monotone transformation of a Jordan region $\bar{r}$ onto $\bar{H}$. For each n let $\bar{H}_n = \bar{D}_1 + \cdots + \bar{D}_n$, let G_n be the continua of $\bar{r}$ carried into $\bar{H}_n$ by $U = U(u)$, and let $\hat{G}_n$ be the set of points of $\bar{r}$ covered by G_n. Now it is easy to see (using the method of T. § 2, Lemma 8) that $\hat{G}_n$ is a continuum and that each point of $\bar{r}$ either belongs to $\hat{G}_n$, is in a bounded complementary domain of a continuum of G_n, or is in a "region" bounded by a portion of a continuum g_n of G_n and a portion of r^* where we include this portion of r^* not in g_n in the "region." Now let G'_n be the collection of continua g'_n, each g'_n being obtained from the corresponding g_n by adding to it all of its complementary domains and all of the "regions" bounded by a portion of g_n and a portion of r^*. Define $U_n(u) = U(u)$ for $u \, \varepsilon \, \hat{G}_n$ and $U_n(u) = U(u_0)$ for u in any of the components of $\bar{r} - \hat{G}_n$, where u_0 belongs to the g_n corresponding to the g'_n containing u. Clearly $U = U_n(u)$ is a monotone transformation of $\bar{r}$ into $\bar{H}_n$ carrying G'_n into $\bar{H}_n$ and $G'_{n,0}$ into $\bar{B}_n$, the base set of $\bar{H}_n$. Let S_n be defined by $x = x_n(u) = X[U_n(u)]$; clearly $\{U_n(u)\}$ converges uniformly to $U(u)$ and S_n converges to S.

Now, by repeated use of Lemma 3 and a simple limit process which merely makes use of the lower semicontinuity of $L(S)$, we see that $L(S_n) \leqq L(S)$ and hence $L(S) = \lim_{n \to \infty} L(S_n)$. By Lemma 2 and Theorem 1, $L(S_n)$ is the sum of the areas of the subsurfaces of S_n (also subsurfaces of S) which correspond to the non-degenerate cyclic elements of $\bar{H}_n$. Hence $L(S)$ is the sum of the areas of the subsurfaces of S corresponding to non-degenerate cyclic elements of $\bar{H}$.

Definition 3. A surface S is said to be *represented generalized conformally* on a double cone $\bar{C}$ by $x = X(U)$ if, when any Jordan region of $\bar{C}$ is mapped conformally (with the obvious conventions at the vertices or equatorial edge) on a plane Jordan region $\bar{r}$ by $U = U(u)$ (this being easily seen to be possible), the function $x(u) = X[U(u)]$ is a generalized conformal vector function, i. e., the components $x^i(u, v)$ are all A. C. T. (Part I, § 2, Definition 1) and the formal expressions E, F, and G satisfy $E = G$, $F = 0$ almost everywhere. It is clear that in such a representation on $\bar{C}$, there is a function M, summable over $\bar{C}$, which is the transform of $E = G$ by the transformation $U = U(u)$, $L(S)$ being given by

$$L(S) = \int\int_{\bar{C}} M \, d\sigma,$$

$d\sigma$ being the element of area on $\bar{C}$. This definition is seen to be consistent for it is easily seen that if $u = u(u')$ is a conformal map of $\bar{r}$ on itself and $x'(u') = x[u(u')]$, then $x'(u')$ is also a generalized conformal vector function on $\bar{r}$.

LEMMA 4. *If $x = X(U)$, $U \varepsilon \bar{C}$, is a generalized conformal representation of S on a double cone $\bar{C}$ and if $U = U(U')$ is a conformal map of $\bar{C}'$ on $\bar{C}$, then $x = X'(U') = X[U(U')]$ is a generalized conformal representation of S on $\bar{C}'$.*

Proof. Let $\bar{R}$ and $\bar{R}'$ be corresponding Jordan regions on $\bar{C}$ and $\bar{C}'$ respectively (the conformal transformation of $\bar{C}'$ into $\bar{C}$ being $1 - 1$ and continuous as is easily seen). It is clear that if we map $\bar{R}$ on a plane Jordan region $\bar{r}$ conformally that the product of this transformation and $U = U(U')$ yields a conformal map of $\bar{R}'$ on $\bar{r}$. Then the lemma follows from Definition 3.

LEMMA 5.[5] *Any polyhedron* Π *which may be mapped non-degenerately on the surface* Σ *of a sphere may be mapped generalized conformally on* Σ, *any three logically distinct points* (Part I, § 2, Definition 4) *of* Π *being made to correspond with three given distinct points of* Σ. *It is easy to see that we may map on a double cone also, for simple formulas give a* $1 - 1$ *conformal* (*with the obvious conventions*) *representation of a double cone on the sphere.*

LEMMA 6. *Let S be a surface of finite Lebesgue area represented non-degenerately on a double cone $\bar{C}$, one end of which is the base set. Then S can be represented generalized conformally on a double cone $\bar{C}_1$ ($= \bar{C}$ perhaps), a given end point being its base set, this point and two others being made to correspond to the first end point and two other points of $\bar{C}$.*

Proof. By Theorem 2, we may choose a sequence $\{\Pi_n\}$ of polyhedra representable non-degenerately on $\bar{C}$ which approach S and are such that $\lim_{n\to\infty} L(\Pi_n) = L(S)$. It is immediate that we may choose a sequence of non-degenerate representations $x = X_n(U)$, $U \varepsilon \bar{C}$, of Π_n on $\bar{C}$ which approach the non-degenerate representation $x = X(U)$ of S on $\bar{C}$. Now let A, B, and C be three distinct points on $\bar{C}$, A say being the base set of $\bar{C}$, and let A', B', and $\bar{C}'$ be three distinct points on $\bar{C}'$, A' being its base set. Now, by Lemma 5 each Π_n may be mapped generalized conformally on $\bar{C}'$, by $x = X'_n(U')$, $U' \varepsilon \bar{C}'$, the points of Π_n corresponding by $x = X_n(U)$ to A, B, and C being mapped

[5] See, for instance, C. Caratheodory, "Conformal representation," *Cambridge Tracts in Mathematics and Mathematical Physics*, No. 28, § 161 and §§ 125-128, particularly.

on A', B', and C' respectively. An argument similar to that of Theorem 1, § 2 (Part I) of this paper shows that we may extract a uniformly convergent subsequence of the $X'_n(U')$ with limit function $X'(U')$, and it is easily seen that $x = X'(U')$ gives a generalized conformal map of S on $\bar{C}'$.

THEOREM 4. *A necessary and sufficient condition that a surface S be of finite area is that there exist a hemicactoid $\bar{H}$ on which S may be represented continuously the representation being generalized conformal on each non-degenerate cyclic element of $\bar{H}$, with M summable over $\bar{H}$. For such representations, $L(S)$ is given by*

$$L(S) = \int\int_{\bar{H}'} M \, d\sigma,$$

$\bar{H}'$ being the sum of the non-degenerate cyclic elements of $\bar{H}$, $d\sigma$ being the element of area on $\bar{H}$.

Proof. The second statement of the theorem and hence the sufficiency of the first statement follows from Lemma 5, § 2 (Part I of this paper), Theorem 2, and Theorem 3.

To show the existence of such a representation of a surface S of finite area, first let S be represented non-degenerately by $x = X'(U')$ on a hemicactoid $\bar{H}' = \bar{B}' + \sum_{m=1} \bar{C}'_m + H'^*$ (T. § 4, Theorem 2), $\bar{B}'$ being its base set and $\bar{C}'_m$ its canonical cactoids. We have $\bar{B}' = \sum_{n=1}^{\infty} \bar{B}'_n + B'$ and $\bar{C}'_m = \sum_{n=1}^{\infty} \bar{C}'_{m,n} + \bar{C}'_m$ where the B' and C'_m are mutually exclusive completely disconnected sets and $\bar{B}'_n$ and $\bar{C}'_{m,n}$ are simple cyclic chains of types B and A respectively joined together as in T. Definition 12, § 2, and T. Theorem 1, § 1.

Let the subsurface corresponding to a non-degenerate cyclic element of one of these chains be represented generalized conformally (since it is of finite area) on that element by $x = \bar{X}'(\bar{U}')$ the points corresponding under $x = X'(U')$ to the end points of the element being mapped on those same end points (Lemma 6). Now on each $\bar{B}'_n$ and $\bar{C}'_{m,n}$ let continuous monotone transformations $T_n^{(2)}$ and $T_{m,n}^{(1)}$ (for $\bar{B}'_n$ and $\bar{C}'_{m,n}$ respectively) be determined as follows: they are to be the identity at all points of their respective chains not in a non-degenerate cyclic element; on a non-degenerate cyclic element let these transformations be one of the monotone transformations $U' = \bar{U}'(U')$, induced by the relation $X'[U'(\bar{U}')] = \bar{X}'(\bar{U}')$, of that element into itself which carries each end point into itself, such a transformation existing by T. § 4, Theorem 3; these $T^i_{k,l}$ are easily seen to be continuous over the whole chain since in each chain, there are only a finite number of non-degenerate cyclic

elements of diameter $\geqq \varepsilon$ for each $\varepsilon > 0$. Now, these transformations satisfy the hypotheses of T. Theorem 5, § 2, so that we can find a hemicactoid $\bar{H}$ with base set $\bar{B}$, etc., such that (1) there are $1-1$ continuous transformations $S^{(1)}_{m,n}$ of $\bar{C}'_{m,n}$ into $\bar{C}_{m,n}$ and $S_n^{(2)}$ of $\bar{B}'_n$ into $\bar{B}_n$ which are conformal (with the obvious conventions at the vertices and edges) between corresponding non-degenerate cyclic elements and (2) the transformations $[S^{(1)}_{m,n}]^{-1}T^{(1)}_{m,n}$ and $[S_n^{(2)}]^{-1}T_n^{(2)}$, defined respectively on $\bar{C}_{m,n}$ and $\bar{B}_n$ unite to form a continuous monotone transformation $U'(U)$ of $\bar{H}$ into $\bar{H}'$ which carries $\bar{B}$ into $\bar{B}'$. If we define $X(U) = X'[U'(U)]$, $U \varepsilon \bar{H}$, $x = X(U)$ is clearly a representation of S on $\bar{H}$, and this representation is generalized conformal on each non-degenerate cyclic element of $\bar{H}$, since for each such element, it is the product of $[S_n^{(2)}]^{-1} \cdot R'$ or $[S^{(1)}_{m,n}]^{-1} \cdot R'$, where R' is the generalized conformal representation $x = \bar{X}'(\bar{U}')$, such a product being generalized conformal by Lemma 4.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

SOME SEPARATION PROPERTIES OF THE PLANE.[1]

By R. E. Basye.

In a previous paper[2] the writer defined the notion of a simply connected set and showed how it was useful in proving a number of separation theorems. Our purpose here is to extend this notion to multiple connectivity and derive some further results for the plane. The abbreviation S. C. will hereafter be used in referring to the paper mentioned above.

Let n be a positive integer. A connected set M is said to be *multiply connected of order n*, or *n-ply connected*, if (1) for each pair of points A and B of M and any relatively closed subset L of M that separates A from B in M there exists a subset of L, consisting of n or less components, which separates A from B in M, and (2) there exists at least one pair of points A and B of M and a relatively closed subset L of M that separates A from B in M such that every subset of L which separates A from B in M has at least n components. This definition becomes the criterion for a connected set to be *n-ply connected in the weak sense* if "separates" is replaced throughout by "weakly disconnects."[3]

Theorem 1. *If, in a metric space, α is a monotonic descending sequence of compact continua which are multiply connected in the weak sense of orders not greater than n, then the product of the sets of α is also multiply connected in the weak sense of order not greater than n.*

The proof is similar to that of Theorem 12 of S. C.

Theorem 2. *If D is a bounded and connected subdomain of the plane whose boundary consists of n mutually exclusive simple closed curves, then $\bar{D}$ is n-ply connected.*

Let F be a closed subset of $\bar{D}$ which separates a point P_1 from a point P_2 in $\bar{D}$. Let F' be a closed subset of F which is the common boundary of two

[1] Presented to the American Mathematical Society, September 7, 1934, under the title, "Multiply connected sets."

[2] "Simply connected sets," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 341-356.

[3] A subset L of a connected set M is said to weakly disconnect a point A from a point B in M if L intersects every connected and relatively closed subset of M which contains $A + B$.

connected domains D_1 and D_2, relative to $\bar{D}$, containing P_1 and P_2 respectively. Then $\bar{D}$ is the sum of two continua, $\bar{D}_1$ and $\bar{D} - D_1$, whose common part is F'. If F' has $n+1$ or more components, it follows from a theorem of S. Straszewicz[4] that $\bar{D}$ has at least $n+1$ complementary domains. But this is not the case. Hence F' has not more than n components.

On the other hand let $J_1, \cdots, J_n$ denote the simple closed curves which constitute the boundary of D. It is known that if Δ is a bounded and connected subdomain of the plane, and b is an arc which lies in Δ except for its endpoints, which lie in the boundary β of Δ, then $\Delta - \Delta \cdot b$ is connected or the sum of two mutually exclusive connected domains Δ_1, Δ_2 according as the endpoints of b lie in different or in the same component of β; in the latter case b lies on the boundary of each of the domains Δ_1, Δ_2. With the aid of this fact there can be constructed n mutually exclusive arcs $b_{12}, b_{23}, \cdots, b_{n-1n}, b_{n1}$ such that b_{kl} lies in D except for its endpoints, which lie in J_k and J_l respectively; and $D - D \cdot \Sigma b_{kl}$ is the sum of two mutually exclusive connected domains d_1, d_2 such that Σb_{kl} lies on the boundary of each. Furthermore it can be seen that $\bar{d}_1 - \Sigma b_{kl}$ and $\bar{d}_2 - \Sigma b_{kl}$ are mutually separated sets. Hence Σb_{kl}, which has n components, irreducibly separates d_1 from d_2 in $\bar{D}$. It follows that $\bar{D}$ is *n-ply* connected.

In passing we remark that, with the aid of Theorem 2 and a generalization of Theorem 7 of S. C., it can be shown that a connected subdomain of the plane is n-ply connected if and only if its complement has exactly $n-1$ bounded components.

THEOREM 3. *Every compact plane continuum which has exactly n complementary domains is multiply connected in the weak sense of order not greater than n.*

Such a continuum can be expressed as the product of a monotonic descending sequence of compact continua each of which is the closure of a connected domain whose boundary consists of n mutually exclusive simple closed curves. Hence Theorem 3 follows from Theorems 1 and 2.

THEOREM 4. *If A and B are two points of a compact plane continuum M having exactly n complementary domains, and G is a countable collection of mutually exclusive closed subsets of M such that G^* is closed and weakly disconnects A from B in M, then there exist n elements or less of G whose sum weakly disconnects A from B in M.*

[4] "Über die Zerschneidung der Ebene durch abgeschlossene Mengen," *Fundamenta Mathematicae*, vol. 7 (1925), p. 173, Theorem II.

By Theorem 3 there exists a closed subset F of G^* which weakly disconnects A from B in M and has not more than n components. No component of F can intersect more than one element of G since, if it did, it would be the sum of countably many (more than one) mutually exclusive closed sets. Hence F is a subset of the sum of n or less elements of G.

THEOREM 5. *Let A and B be two points of a connected and locally arcwise connected metric space S, and let G be a countable collection of closed sets such that* (1) *the common part of every pair of elements of G is the closed set H (which may be vacuous),* (2) *if $b_1, \cdots, b_{n+1}$ are $n+1$ arcs from A to B that lie in $S-H$, then $b_1 + \cdots + b_{n+1}$ lies in a compact set which is multiply connected in the weak sense of order n or less and whose closure contains no point of H, and* (3) *G^* is locally compact. If G^* separates A from B in S then G contains n or less elements whose sum separates A from B in S.*

The proof is similar to that of Theorem 3 of S. C.

LEMMA Q. *In a sphere S let G be a countable collection of closed sets such that the common part of every pair of elements of G is N, a closed set having exactly n components, where n is a positive integer. If G^* separates a point A from a point B in S, then G contains n or less elements whose sum separates A from B in S.*

Consider any $n+1$ arcs from A to B which have no point in common with N, and let β denote their sum. If $N_1, \cdots, N_n$ denote the components of N, let $J_1, \cdots, J_n$ denote n mutually exclusive simple closed curves such that $J_i (i = 1, \cdots, n)$ separates β from N_i in S. The boundary of the complementary domain Δ of $J_1 + \cdots + J_n$ which contains β consists of one or more of the curves J_i; hence $\bar{\Delta}$ is homeomorphic to the closure of a bounded and connected plane domain whose boundary consists of n or less mutually exclusive simple closed curves. Hence, by Theorem 2, $\bar{\Delta}$ is multiply connected of order not greater than n. Therefore, by Theorem 5, G contains n or less elements whose sum separates A from B in S.

With the aid of the preceding lemma we can generalize, as follows, a theorem of R. L. Moore.[5]

THEOREM 6. *In a plane S let G be a countable collection of closed sets such that* (1) *the common part of every pair of elements of G is N, a closed*

[5] "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, vol. 13, p. 298, Theorem 113.

set having exactly n components, where n is a positive integer, and (2) *there exists not more than one element g of G such that g — N is not compact. If G^* separates a point A from a point B in S, then G contains n or less elements whose sum separates A from B in S.*

Let Σ be a sphere, P a point of Σ, and H a homeomorphism of S into $\Sigma - P$. Let G', N', A', B' be the images of G, N, A, B, respectively, under H. Let G'' be the collection whose elements are defined as follows. Let g' denote an element of G'. If P is not a limit point of g' then g' is an element of G'', and if P is a limit point of g' then $g' + P$ is an element of G''. If P is not a limit point of any element of G' then $N' + P$ is an element of G''. The elements of G'' are closed in Σ and the common part of each pair is the same set, which we call N''. Now N'' does not have more than n components. Hence, since G''^* separates A' from B' in Σ, there exists, by Lemma Q, n or less elements of G'' whose sum ν separates A' from B' in Σ. Thus $\nu - \nu \cdot P$ separates A' from B' in $\Sigma - P$ and moreover is the sum of not more than n elements of G'. Therefore the image of $\nu - \nu \cdot P$ under H separates A from B in S and is the sum of not more than n elements of G.

In Theorem 21 of S. C. the writer has given a certain extension of a result [6] of Rutt and Roberts. With the aid of Theorem 6 we can extend this result in another direction, as follows.

THEOREM 7. *In a plane S let N be a closed set having exactly n components, where n is a positive integer. Let G be any collection of closed sets and H a countable subcollection* (*possibly vacuous*) *of G such that* (1) *the common part of each pair of elements of G is N,* (2) *every component of every element of G — H intersects N,* (3) *there exists not more than one element g of G such that g — N is not compact, and* (4) *G^* is closed. If G^* separates a point A from a point B in S then G contains n or less elements whose sum separates A from B in S.*

Suppose the contrary. Let G' denote the collection of all elements each of which is either an element of H or a component of the sum of N and a component of a set obtained by subtracting N from an element of $G - H$. Let H' denote the collection whose elements are those of H and that element g of G, if there exists one, such that $g - g \cdot N$ is not compact. There exists a subset F of G'^* which is closed, separates A from B in S, contains every

[6] N. E. Rutt, "On certain types of plane continua," *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 815, Theorem IV and Corollary IV; and J. H. Roberts, "Concerning collections of continua not all bounded," *American Journal of Mathematics*, vol. 52 (1930), p. 553, Theorem I.

element of G' with which it has a point of $G'^* - (H^* + N)$ in common, and is irreducible with respect to these three properties. The set F is a continuum.

With the aid of Theorem 6 there can be constructed $4n$ arcs $b_1, \cdots, b_{4n}$ from A to B which have no points in common with $H'^* + N$ and such that, if P_i $(i = 1, \cdots, 4n)$ denotes the first point of b_i in the order from A to B which lies in F, the points P_i all lie in different elements of G'. The element g'_i of G' which contains P_i contains a component of N. Hence there exists a component of N which is common to four of the elements g'_i, say g'_1, g'_2, g'_3, g'_4. By an argument similar to one used in Theorem 21 of S. C. it can be shown that two of these four continua, say g'_1 and g'_3, have the property that $g'_1 + g'_3$ separates $g'_2 - g'_2 \cdot N$ from $g'_4 - g'_4 \cdot N$ in F. Thus $F - (g'_1 + g'_3) = F_2 + F_4$, where F_2 and F_4 are mutually separated sets containing P_2 and P_4 respectively. Consider the two sets $R_2 = F_2 + (g'_1 + g'_3)$ and $R_4 = F_4 + (g'_1 + g'_3)$. The common part of R_2 and R_4 is a continuum and their sum is F. Hence either R_2 or R_4, say the former, separates A from B in S. But R_2 is a proper closed subset of F which contains every element of G' with which it has a point of $G'^* - (H^* + N)$ in common. This is contrary to the construction of F.

If the set N is not restricted as to the number of its components, the following weaker conclusion will hold true.

THEOREM 8. *In a plane S let N be a non-vacuous closed set, G any collection of closed sets, and H a countable subcollection (possibly vacuous) of G such that* (1) *the common part of each pair of elements of G is N,* (2) *every component of every element of $G - H$* intersects *N,* and (3) *G^* is closed. If G^* separates a point A from a point B in S, then G contains a countable subcollection K such that K^* intersects every compact continuum which contains $A + B$.*

Suppose that G^* is compact. Let $\Delta_1, \Delta_2, \cdots$ be a sequence of domains closing down on N such that the closure of each domain has not more than a finite number of components and contains neither A nor B. Consider the collection G_i $(i = 1, 2, \cdots)$ of all elements each of which is the sum of $\bar{\Delta}_i$ and an element of G. This collection satisfies the conditions of Theorem 7 and hence contains a finite number of elements whose sum L_i separates A from B in S. Denote by M_i the finite collection of those elements of G which intersect $L_i - \bar{\Delta}_i$. We take for K the collection of all elements of G which are elements of some M_i. Let C be a continuum which contains $A + B$. If C intersects N it intersects every element of K. If C does not intersect N there exists an integer r such that $\bar{\Delta}_r$ contains no point of C; therefore C contains a point of $L_r - \bar{\Delta}_r$ and hence a point of an element of K.

The case where G^* is not compact can be reduced by an inversion of the plane to the one considered.

It has been proved by Kuratowski [7] that if three compact plane continua have a point in common, and their sum separates a point A from a point B in the plane, then there exists a pair of these continua whose sum separates A from B in the plane. The writer has elsewhere [8] obtained a generalization of this result for which the containing space is any subcontinuum of the plane. An extension in another direction, in which the number of continua is not restricted, will here be given.

THEOREM 9. *In a plane S let N be a non-vacuous closed set and let G be any collection of continua such that* (1) *no two elements of G have a point of $S - N$ in common,* (2) *N contains r points* (r *finite*) *whose sum intersects every element of G,* (3) *$G^* + N$ is closed, and* (4) *the collection G is upper semicontinuous. If $G^* + N$ separates a point A from a point B in S, there exists a subcollection H of G, containing not more than $2r$ elements, such that $H^* + N$ separates A from B in S.*

Suppose first that $G^* + N$ is compact. Let $\{\Delta_i\}$ be a sequence of compact domains closing down on N such that $\bar{\Delta}_i$ $(i = 1, 2, \cdots)$ contains neither A nor B and has only a finite number of components. Consider the collection G_i $(i = 1, 2, \cdots)$ of all elements each of which is the sum of $\bar{\Delta}_i$ and an element of G. By Theorem 7 there exists a finite number of elements of G_i whose sum separates A from B in S. Therefore, by a theorem [9] of the writer, there exist $2r$ or less elements of G_i whose sum separates A from B in S. Denote these elements by $g_i^1, \cdots, g_i^{2r}$. There exists a subsequence $\{i_n\}$ $(n = 1, 2, \cdots)$ of the positive integers such that the sequence $\{g_{i_n}^j\}$ $(j = 1, \cdots, 2r)$ has a sequential limiting set R^j. With the aid of condition (4) of the hypothesis it follows that R^j is a subset of $N + \nu^j$, where ν^j denotes some element of G. Hence $R = (R^1 + \cdots + R^{2r})$ is a subset of $N + (\nu^1 + \cdots + \nu^{2r})$. But R separates A from B in S. Hence for H we may take the collection $\nu^1, \cdots, \nu^{2r}$.

If $G^* + N$ is not compact, the proof can be made by an inversion of the plane.

Theorem 9 is not true if condition (4) is omitted.

THE UNIVERSITY OF TEXAS.

[7] "Théorème sur trois continus," *Monatshefte für Mathematik und Physik*, vol. 36 (1929), pp. 77-80.

[8] "Concerning two internal properties of plane continua," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 670-674.

[9] *Loc. cit.*

THE ANALYSIS OF CERTAIN CURVES BY MEANS OF DERIVED LOCAL SEPARATING POINTS.

By G. E. SCHWEIGERT.

1. *Introduction.* It is well known[1] that the removal of the local separating points[2] from a regular curve leaves a set of dimension zero. This suggests that a similar result may be true for an hereditarily locally connected continuum, that is, a continuum having the property that every subcontinuum is locally connected. However an example by Gehman[3] shows that non-degenerate components may remain in such a continuum after the removal of its local separating points. In this example a line segment is the only component that remains and if we remove in addition the local separating points of the line segment, the result is a set of dimension zero. Thus although it is clearly necessary to remove points other than the local separating points, the situation in this example suggests that *given an hereditarily locally connected continuum, a set of dimension zero may be reached by a countable iteration of the process of removing the local separating points from the components that remain at each successsive stage.* It is to this problem that sections 2, 3, 4, and 5 of this paper are devoted.

2. *Derived local separating points.* For any subset M of a compact, metric, hereditarily locally connected continuum H, let $L_0(M)$ denote the set of all local separating points of M. Then we may write $H - L_0(H) = \Sigma C_k{}^1 + P_1$ where the $C_k{}^1$ are the non-degenerate components in the set on the right and P_1 is a set of dimension zero. In doing this we make use of the fact[4] that an arbitrary subset of an hereditarily locally connected continuum is made up of non-degenerate components and a zero-dimensional set, with the former in the form of a (null) sequence. Now let $L_1(H) = \Sigma L_0(C_k{}^1)$, and proceed to represent the set $H - [L_0(H) + L_1(H)]$ in a similar fashion. This process

[1] G. T. Whyburn, "Über die Strukture regulärer Kurven," *Wiener Akademie Anzeiger* (1930), Nr. 6.

[2] If p is a cut-point of a connected neighborhood R in a connected and locally connected space then p is said to be a *local separating point*, that is, $R - p = R_1 + R_2$ where $\bar{R}_1 R_2 = 0 = R_1 \bar{R}_2$ and neither R_1 nor R_2 is vacuous.

[3] *Annals of Mathematics*, vol. 27 (1926), p. 43.

[4] G. T. Whyburn, "Concerning hereditarily locally connected continua," *American Journal of Mathematics*, vol. 53, no. 2 (1931), Corollary c, p. 379 and Theorem 4, p. 377.

continues under an inductive scheme of definition which uses the transfinite ordinal numbers. For this purpose we assume that the components C_N^γ have been defined as the components of the set $H - \sum_{\beta<\gamma} L_\beta(H)$ for every $\gamma < \alpha$, and also that $L_\gamma(H) = \Sigma L_0(C_N^\gamma)$ for every $\gamma < \alpha$. Then the general terms are $H - \sum_{\beta<\alpha} L_\beta(H) = \Sigma C_i^\alpha + P_\alpha$ and $L_\alpha(H) = \Sigma L_0(C_i^\alpha)$ where: (a) for each pair (i,α), $i = 1, 2, 3, \cdots$ and α transfinite, C_i^α is a component of $H - \sum_{\beta<\alpha} L_\beta(H)$; (b) P_α is a set of dimension zero.

For each i the set C_i^α is said to be a *derived component of index* α relative to H. Similarly the points in the set $L_\alpha(H)$ are the *derived local separating points of index* α relative to the continuum H. Under this notation H is denoted by C_1^0. Due to a theorem by Wilder [5] it is known that each derived component is a locally connected set.

If we make use of the fact that the derived local separating points are dense in each derived component [6] we may put the solution to our first problem in the form of the next proposition.

THEOREM. *If H is any compact, metric, hereditarily locally connected continuum there exists a number α of the first or second number class such that $L_\alpha(H) = 0$.*

If $\alpha(H)$ is the first such ordinal number α, then $\alpha(H)$ is said to be the *index* of the continuum H.

The above theorem [7] will be established with the aid of the work of Menger and Reschovsky on rational curves.

3. *Unilateral and bilateral limit points.* A generic member of the derived components of index α will be denoted by C^α. Whenever C^α, C^β, C^γ, etc., are used they are to be thought of as members of a monotone sequence which terminates with the component of highest index.

The symbol R denotes any fixed neighborhood in H having a countable boundary and of such a nature that it is in diameter less than that of any component with which it appears. The boundary of any such neighborhood is

---

[5] R. L. Wilder, *Proceedings of the National Academy of Sciences*, vol. 15 (1929), p. 616.

[6] This result is an easy consequence of the fact that any isolated point of an irreducible cutting between two points of a connected and locally connected set M is a local separating point of M. For the theorem quoted here see G. T. Whyburn, *Monatshefte für Mathematik und Physik*, vol. 36 (1929).

[7] A more elementary proof using only the material now at hand together with the fact that H is a rational curve could be given at this point.

a countable cutting of H for which we consider only the particular separation $H - F(R) = R + (H - \bar{R})$. Similarly for the cutting $F(R)C^\alpha$ of C^α we use but one separation $C^\alpha - F(R)C^\alpha = C^\alpha R + C^\alpha(H - \bar{R})$.

If X is a cutting of a set M and the point x belongs to X, then x is said to be a *unilateral limit point of M with respect to the fixed separation* $M - X = M_1 + M_2$ provided x is a limit point of M_1 or M_2 but not of both these sets. If x is a limit point of both M_1 and M_2 then x is a *bilateral limit point of M with respect to the separation.* Since we are to be concerned only with cases in which there is a known separation (as in the paragraph above) the shorter terms unilateral and bilateral limit point *of M* will be used.

It can easily be shown that if x is a unilateral limit point of C^α then x is a unilateral limit point of C^β provided that $x \epsilon F(R)C^\beta$ and $\beta > \alpha$. In regard to the bilateral limit points, what amounts to the converse is true, namely, if x is a bilateral limit point of C^α then x is a bilateral limit point of any component of smaller index.

The α-th derivative of the boundary of any neighborhood N will be denoted by $F(N)^\alpha$. A set E belonging to the cutting $F(R)C^\alpha$ of C^α will be said to be *of type B^α* provided there exists for each point $y \epsilon E$ a number $\alpha(y) < \alpha$ such that y is a unilateral limit point of $C^{\alpha(y)}$. It follows that y is a unilateral limit point of all those derived components of index greater than $\alpha(y)$, including C^α. Another most useful result in this connection is that y is a unilateral limit point of $C^{\alpha-1}$ when α is an isolated number. If E is a set of type B^α we shall write $E = B^\alpha$. The only sets E to be used here are the particular type shown in the next paragraph.

We now prove a lemma from which the principal proposition of this paper is to be deduced. The lemma is stated in terms of the symbols just developed and its second part is a direct consequence of the first.

LEMMA. (1) $[F(R) - F(R)^\zeta]C^\zeta = B^\zeta$; (2) *If x is a bilateral limit point of C^ζ and $x \epsilon [F(R) - F(R)^{\zeta+1}]C^\zeta$, then x is a derived local separating point of index ζ.*

Proof. When $\zeta = 1$ it must be shown in part 1 that there is *no* bilateral limit point x of C^0 (that is, of H) among those isolated points of $F(R)$ which are in C^1. In this connection we recall that the set $F(R)$ is a *countable* cutting of H. This allows us to obtain a separation of H by means of its bilateral limit points alone and from this stage it is but an easy step to show that x is a local separating point of H. Thus x is not a point of C^1 which is contrary to the hypothesis for the existence of that point.

The point x described in part 2 must lie in $F(R)^1 - F(R)^2$ because

it is a bilateral limit point of C^1 and therefore cannot be an isolated point in $F(R)$. But this means that x is an isolated point of X_0, where X_0 is the set of all bilateral limit points of C^1. Then by means of the same argument as was used above it follows that x is a point of $L_1(H)$.

Now for the purpose of induction we assume that the Lemma is true for $\zeta < \eta$ and show that it holds for $\zeta = \eta$.

If we assume the contrary of what is to be shown, we have that some point x is a bilateral limit point of C^ζ for all $\zeta < \eta$ where $x \epsilon [F(R) - F(R)^\eta]C^\eta$. It follows that if η is a limit number, there exists a number $\lambda < \eta$ such that x non-$\epsilon F(R)^\lambda$. But then $x \epsilon [F(R) - F(R)^\lambda]C^\lambda$, and furthermore x is a bilateral limit point of C^λ because that property holds for C^ζ where $\zeta < \eta$. This is impossible in view of the fact that $[F(R) - F(R)^\lambda]C^\lambda = B^\lambda$ holds by assumption and hence no point of the set on the left is a bilateral limit point of C^λ. If on the other hand η is an isolated number, we know that x is a bilateral limit point of $C^{\eta-1}$. Moreover $x \epsilon [F(R) - F(R)^\eta]C^{\eta-1}$, since $x \epsilon C^\eta$. Then by part 2 for $\zeta = \eta - 1$, we find that $x \epsilon L_{\eta-1}(H)$. This is impossible when $x \epsilon C^\eta$.

It remains to be shown that if x is a bilateral limit point of C^η and $x \epsilon [F(R) - F(R)^{\eta+1}]C^\eta$, then $x \epsilon L_\eta(H)$. Certainly such a point x cannot belong to $F(R) - F(R)^\eta$ because of the results just above; hence $x \epsilon F(R)^\eta - F(R)^{\eta+1}$. Now let X_0 be the set of all bilateral limit points of C^η. Then $X_0[F(R) - F(R)^\eta]C^\eta = 0$ because the last two factors form a set of type B^η. It follows that x is an isolated point of X_0 and hence by the argument used earlier that $x \epsilon L_\eta(H)$. This completes the induction, that is, the proof of the Lemma.

4. *Relation between index and genus.* A point p of a set M is said to be of *genus α in M* [8] provided: (a) there exists arbitrarily small neighborhoods in M closing down on p in such a way that the α-th derivative of the boundary of each neighborhood is vacuous; (b) for some $\epsilon > 0$ and every $\beta < \alpha$ there exists *no* neighborhood in M which contains p, is of diameter less than ϵ and has the further property that the β-th derivative of the boundary is vacuous. The genus of the point p will be denoted by $g(p)$.

Suppose there is a point p in a derived component C^α such that $g(p) \leq \alpha$. Then since p is also in $C^{g(p)}$, there exists a neighborhood R in H such that R contains p and does not contain all points of $C^{g(p)}$. Moreover R can be so chosen that $F(R)^{g(p)} = 0$ and $F(R)$ is countable. It follows from the fact that $C^{g(p)}$ is connected that there exists a point x of $F(R)C^{g(p)}$ such that x is a bilateral limit point of $C^{g(p)}$. Thus $x \epsilon [F(R) - F(R)^{g(p)}]C^{g(p)}$ because

[8] *Geschlecht*—definition due to Menger. See *Kurventheorie* (Teubner), p. 294.

$F(R)^{g(p)} = 0$. This is contrary to the Lemma in section 3, hence we have shown that the following proposition is true.

THEOREM. *If H is any compact, metric, hereditarily locally connected continuum and C^{α} is any one of its derived components of index α, then each point of C^{α} is of genus greater than α.*

5. *Solution to the problem stated in section* 1. It has been shown by Reschovsky[9] that there is associated with each rational curve K a number α of the first or second class such that the genus of each point of K is less than or equal to α. The smallest number $g(K)$ for which this statement holds is called the genus of the rational curve.[10]

This classification applies to the hereditarily locally connected continuum H for it has been established by G. T. Whyburn[11] that each such continuum is a rational curve. Thus it is an obvious corollary to the Theorem immediately above that $g(H) \geq \alpha(H)$. This means that $\alpha(H)$ is also of the first or second number class, consequently the Theorem stated in section 2 is proven and thus we have a positive solution to our original problem.

6. *Regular and rational bases.* A subset B of an arbitrary set M is said to be a *regular basis* for M provided that, (1) each regular point of M is contained in arbitrarily small neighborhoods N_ϵ in M which close down on p in such a way that $F(N_\epsilon)$ is finite and is contained in B, (2) when p is of order k, the neighborhoods N_ϵ can be chosen so that $F(N_\epsilon)$ consists of exactly k points of B. Similarly B is said to be a *rational basis* for M provided that each rational point of M is contained in arbitrarily small neighborhoods in M which close down on that point and have countable boundaries in B. Obviously a set of dimension zero remains on removing a rational (regular) basis from a rational (regular) curve.

In the paper by G. T. Whyburn referred to in the introduction it is shown that B will be a countable regular basis for a regular curve K provided that $B = T + D$, where T is the set of all local separating points of order greater than 2 in K and D is a countable set which is dense on each free arc in K. It is our purpose to establish a proposition which is analogous to this one in that it enables us to choose a rational basis for a curve H from its derived local separating points. The next Theorem is offered as a solution to this problem.

THEOREM. *Let H be any compact, metric, hereditarily locally connected*

[9] Reschovsky, *Fundamenta Mathematica*, vol. 15 (1930), p. 18.
[10] *Loc. cit.*
[11] "Concerning hereditarily locally connected continua."

continuum and let B be any subset of H such that $B = \sum_{i,a} T_i^a + D$, where D is a countable set, and such that in regard to each derived component C_i^a: (a) *T_i^a is the set of all derived local separating points of order greater than two in that component and of index equal to that of the component,* (b) *D is dense on each free arc. Then B is a countable rational basis for H and the part of B which is contained in each derived component is a countable regular basis for that component.*

In the statement of this Theorem we again follow the convention that H is uniquely the derived component of index zero; hence when H is a regular curve, this Theorem reduces to the one quoted in the above paragraph. The proof of this Theorem is an almost direct result of the following Lemma.

LEMMA. *If C is any connected and locally connected set and B is any subset of C such that $B = T + D$, where T is the set of all local sepaarting points of order greater than two in C and D is a countable set which is dense on each free arc in C, then B is a regular basis for C.* Moreover B is a *countable* regular basis provided the nodules in C form a null sequence.

This Lemma follows from an argument based on the work of G. T. Whyburn.[12]

Since each derived component C_i^a of the curve H is locally connected we may apply the Lemma to each of these components and obtain a regular basis $B_i^a = T_i^a + D_i^a$. This basis is a countable set in each case because the nodules are certainly in a null sequence when the containing space is hereditarily locally connected. Let T and D denote the sets $\sum_{i,a} T_i^a$ and $\sum_{i,a} D_i^a$ respectively and denote their sum by B. From this approach it is obvious that the part of B that is contained in each derived component is a regular basis for that component. Moreover there are but a countable number of derived components hence B is a countable set. Thus in order to complete the proof of the Theorem we have to show that B is a rational basis.

It is first shown that the non-degenerate components of $C_i^a - B_i^a$ are exactly the components of $C_i^a - L_0(C_i^a) = \Sigma C_k^{a+1} + P_{a+1}$ after which it is immediate that $H - B$ is a set of dimension zero. For the purpose of comparing the two sets, let $C_i^a - B_i^a = \Sigma \Gamma_j^{a+1} + Q_{a+1}$ where the Γ_j^{a+1} are the non-degenerate components of the containing set and Q_{a+1} is a set of dimension zero. Now assume there exists a point p in some fixed component Γ_j^{a+1} such that p non-ϵ C_k^{a+1} for all k. Then p is either a point of $L_0(C_i^a) - B_i^a$ or a point of P_{a+1}. When p is in the former set, let $q \epsilon \Gamma_j^{a+1}$ be a point distinct

---

[12] See his papers "On the structure of connected and connected im kleinem point sets," *Transactions of the American Mathematical Society*, vol. 32 (1930), pp. 926-943, and "Über die Struckture regulärer Kurven," *Wiener Anzeiger* (1930), Nr. 6.

from p. Then there exists a neighborhood R in C_i^a such that R contains p but does not contain q and such that $F(R) = x, y \epsilon B_i^a$. (The point p is of order two by hypothesis.) However Γ_j^{a+1} is connected and consequently $F(R)$ contains a point z which is distinct from x and y. This is a contradiction. When p is in the latter set P_{a+1}, the point p is contained in arbitrarily small neighborhoods in $C_i^a - L_0(C_i^a)$ each of which does not contain all of that set and has a vacuous boundary. On the other hand we have just shown that Γ_j^{a+1} can contain no point of $L_0(C_i^a) - B_i^a$ hence $\Gamma_j^{a+1} - p$ is contained in a component C_k^{a+1} for some k. These two situations are inconsistent and bring us to the result that $H - B$ is a set of dimension zero. But it obviously follows that $(H - B) + p$, where p is any point of B, is also a set of dimension zero and consequently B is a rational basis. This completes the proof of the Theorem.

It is to be noted that in each case where a basis is mentioned in this section, that basis is prescribed, except for a certain freedom in the choice of the sets which are dense on the free arcs, by the set which contains it. Thus it may be said that each basis is a *natural* basis.

7. *Conclusion.* We are now in a position to demonstrate that any hereditarily locally connected continuum H is topologically contained in one which is of index 1. For this purpose we need the fact that there exists a locally connected Universal Curve in Three-space which contains a subset topologically equivalent to each compact, metric curve.[13] Now let H be any hereditarily locally connected continuum and let T be a homeomorphism throwing H into a subset $T(H)$ of the Universal Curve. Let $T(B) = b_1 + b_2 + b_3 + \cdots$ be the image under T of a countable rational basis B in H. Then add to $T(H)$ a line segment $L(1)$ of a length 1 having the point b_1 in common with $T(H)$. In general $L(n)$ is to be a member of a monotone sequence of disjoint line segments which meets $T(H)$ in just one point $b_n \epsilon T(B)$ and has the length $1/n$. The curve $T(H) + \Sigma L(n)$ will then be of index 1.

One can easily see that if the original curve H is contained in a space of a finite number of dimensions, it can be augmented in the manner described above and the resulting curve will have the same genus as the original curve. This suggests to the author that, while some examples of his (similar to the one by Gehman) showing plane curves with indices $2, 3, 4, \cdots, \omega$ are cases for which $\alpha(H) = g(H)$, this method will alter those examples to show that there are cases in which $\alpha(H)$ takes the full range allowed by the inequality $\alpha(H) \leq g(H)$. In contrast to this we know that if H is a regular curve $\alpha(H) = g(H) = 1$.

THE JOHNS HOPKINS UNIVERSITY.

[13] Menger, *Kurventheorie*, p. 365.

ON THE INDEPENDENCE OF HILBERT AND ACKERMANN'S POSTULATES FOR THE CALCULUS OF PROPOSITIONAL FUNCTIONS.

J. C. C. McKinsey.

Hilbert and Ackermann have given a set of postulates for the calculus of propositional functions,[1] but they state [2] that the independence of the set has not been investigated. In this paper I show that the postulates in question are independent.

I first set forth the undefined marks and the postulates of the system. The undefined marks [3] are:

1) Propositional-variables: $X, Y, Z, \cdots$.

2) Individual-variables: $x, y, z, \cdots$.

3) Function-variables: $F(\ ), G(\ ), \cdots$.

4) The signs (x), (y), etc. These can occur before a function-variable into which has been substituted an individual-variable: thus, $(x)F(x)$, etc.

5) The signs (Ex), (Ey), etc. Giving $(Ex)F(x)$, etc.

6) The sign $\vee$.

7) The sign $\overline{}$.

In the expression $(\alpha)\beta$ the sign (α) is called the *universal-operator*; and in $(E\alpha)\beta$, the sign $(E\alpha)$ is called the *existential-operator*. Universal-operators and existential-operators are together referred to as *operators*. In both operators the α is always an individual-variable, and it is termed the *operator-variable*. β is called the *scope* of α. If an individual-variable x occurs within the scope of an operator having x as operator-variable, then x is called an *apparent* variable, or a *bound* variable. Other individual-variables are called *real* variables, or *free* variables.

---

[1] D. Hilbert and W. Ackermann, *Grundzüge der Theoretischen Logik*, Julius Springer, 1928, pp. 53-54.

[2] *Loc. cit.*, p. 68. Recently, however, a set of independence examples for a closely related set of postulates has been given by B. Notcutt, "A set of independent postulates for propositional functions of one variable," *Annals of Mathematics*, vol. 37 (1935), pp. 670-678. The similarity of Notcutt's postulates to Hilbert and Ackermann's arises from the fact that both sets are modifications of the set given in *Principia Mathematica*.

[3] The undefined terms are not given by Hilbert and Ackermann in the usual mathematical form; that is to say, the authors speak simply of undefined "marks," instead of undefined "classes," "operations," and "relations."

The term "expression" (as used in the postulates) is precisely defined as follows:

1. A sign for a proposition (as X, Y, etc.) is an expression.
2. A sign for a propositional function in which each blank space is filled either by an individual-variable, or by the proper name of an individual is an expression.
3. If A is an expression, so is $\bar{A}$; and, in case the free variable x occurs in A, so is $(x)A$ and $(Ex)A$.
4. If A and B are expressions, so is $A \vee B$.
5. But the restriction is to be understood that in an expression we never have overlapping scopes of a like-indicated variable. Thus, for example, $(x)F(x) \vee (Ex)G(x)$ is an expression, but $(x)(y)[F(x) \vee (Ex)G(x, y)]$ is not.

The complex sign $A \rightarrow B$ is defined to be $\bar{A} \vee B$. From 3 and 4, it is seen that $A \rightarrow B$ is an expression if A and B are expressions.

On these undefined marks the following postulates [4] are imposed:

a) $X \vee X \rightarrow X$.
b) $X \rightarrow X \vee Y$.
c) $X \vee Y \rightarrow Y \vee X$.
d) $(X \rightarrow Y) \rightarrow [Z \vee X \rightarrow Z \vee Y]$.
e) $(x)F(x) \rightarrow F(y)$.
f) $F(y) \rightarrow (Ex)F(x)$.

α) (*Rule of Substitution*). If in any correct formula we substitute any expression for a propositional-variable (making the same substitution, however, wherever this one propositional-variable occurs) then the resulting formula is also correct; similarly, for a function-variable with arguments $x, y, \cdots, u$ we may substitute an expression which depends on $x, y, \cdots, u$ (or, if desired, on still more variables). An individual-variable may be replaced by an individual-variable denoted by a different letter, or by a proper name from the range of the variable.[5]

[4] Postulate a) may be understood as reading "$X \vee X \rightarrow X$ is a correct formula"; and similarly for b)–f).

[5] Postulate α) is followed by the remark that it is to be understood that, when a substitution is made into an expression within the scope of an apparent variable, the expression substituted should not contain the apparent variable in question. Thus, for example, we cannot go from the correct formula $Z \vee (x)F(x) \rightarrow (x)[Z \vee F(x)]$ to the formula $G(x) \vee (x)F(x) \rightarrow (x)[G(x) \vee F(x)]$.

β) (*Rule of Inference*). If A is a correct formula and $A \rightarrow B$ is a correct formula, then B is a correct formula.

γ) Let $B(x)$ be any expression which depends on x, and let A be any expression which does not depend on x. Then if $A \rightarrow B(x)$ is a correct formula, $A \rightarrow (x)B(x)$ is also a correct formula. And if $B(x) \rightarrow A$ is a correct formula, $(Ex)B(x) \rightarrow A$ is also a correct formula.

I here add certain lemmas which will be used in the independence proofs.

LEMMA I. *The following formula is deducible from Hilbert's postulates:*

$$F(y) \rightarrow F(y).$$

Proof.

(1)	$\left[\text{d}), \frac{X \vee X}{X}, \frac{X}{Y}, \frac{\bar{X}}{Z}\right]$	$(X \vee X \rightarrow X) \rightarrow [\bar{X} \vee (X \vee X) \rightarrow (\bar{X} \vee X)]$
(2)	[a)]	$X \vee X \rightarrow X$
(3)	[(1), (2), β)]	$\bar{X} \vee (X \vee X) \rightarrow (\bar{X} \vee X)$
(4)	[(3), Def. of $\rightarrow$]	$(X \rightarrow X \vee X) \rightarrow (X \rightarrow X)$
(5)	$\left[\text{b}), \frac{X}{Y}\right]$	$X \rightarrow X \vee X$
(6)	[(4), (5), β)]	$X \rightarrow X$
	$\left[(6), \frac{F(y)}{X}\right]$	$F(y) \rightarrow F(y).$

LEMMA IIa. *If K_1 is independent of x, and $K_2(x)$ depends on x, and if $K_1 \rightarrow K_2(x)$ is deducible from Hilbert's postulates, then $K_1 \rightarrow (Ex)K_2(x)$ is also deducible from Hilbert's postulates.*

LEMMA IIb. *Similarly, if $K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates, then $(x)K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates.*

Proof. I first show that if $A \rightarrow B$ and $B \rightarrow C$ are deducible from Hilbert's postulates, where A, B, and C are any expressions, then $A \rightarrow C$ is deducible also. For substituting B for X, C for Y, and $\bar{A}$ for Z in postulate d), and applying the definition of $\rightarrow$, we get

$$(B \rightarrow C) \rightarrow [(A \rightarrow B) \rightarrow (A \rightarrow C)].$$

Then since $B \rightarrow C$ is a correct formula, we have, by postulate β)

$$(A \rightarrow B) \rightarrow (A \rightarrow C).$$

Then since $A \rightarrow B$ is a correct formula, we have, again by postulate β), that $A \rightarrow C$ is a correct formula.

Now suppose we have $K_1 \rightarrow K_2(x)$ given as being deducible from Hilbert's postulates. Substituting $K_2(x)$ for $F(x)$ in postulate f), we have

$K_2(x) \rightarrow (Ex)K_2(x)$. Hence by the first part of the proof, the formula $K_1 \rightarrow (Ex)K_2(x)$ is deducible from Hilbert's postulates.

Similarly, if we have $K_2(x) \rightarrow K_1$, then, since by postulate e) we also have $(x)K_2(x) \rightarrow K_2(x)$, it follows that $(x)K_2(x) \rightarrow K_1$ is deducible from Hilbert's postulates.

LEMMA III. *The following formulas are not deducible from Hilbert's postulates:*

$$\text{a) } (Ex)F(x) \rightarrow F(y)$$
$$\text{b) } F(y) \rightarrow (x)F(x).$$

Proof. The following example is given by Hilbert in another connection.[6] A formula is said to be satisfied if by the following process it is transformed into a correct formula of the calculus of propositions: First generalize all the free variables by putting the corresponding universal-operators at the left of the formula. Then[7] if $A(x)$ is any expressions, substitute $A(0)$ & $A(1)$ for $(x)A(x)$, and substitute $A(0) \vee A(1)$ for $(Ex)A(x)$. Continue this process until all apparent variables have been eliminated. The formula will then contain only propositional variables and symbols of the form $F(0)$ and $F(1)$. Then substitute for $F(0)$, wherever it occurs, the same propositional variable, for $F(1)$ another, and so on. By this process, postulate e), for example, becomes successively:

$$(y)[(x)F(x) \rightarrow F(y)]$$
$$[(x)F(x) \rightarrow F(0)] \,\&\, [(x)F(x) \rightarrow F(1)]$$
$$[F(0) \,\&\, F(1) \rightarrow F(0)] \,\&\, [F(0) \,\&\, F(1) \rightarrow F(1)]$$
$$[X \,\&\, Y \rightarrow X] \,\&\, [X \,\&\, Y \rightarrow Y].$$

The last formula is a correct formula of the calculus of propositions. Thus postulate e) holds for this interpretation. In a similar way it may be shown that all the other postulates are satisfied by this interpretation.

Consider now the formula $(Ex)F(x) \rightarrow F(y)$. This becomes successively:

$$(y)[(Ex)F(x) \rightarrow F(y)]$$
$$[(Ex)F(x) \rightarrow F(0)] \,\&\, [(Ex)F(x) \rightarrow F(1)]$$
$$[F(0) \vee F(1) \rightarrow F(0)] \,\&\, [F(0) \vee F(1) \rightarrow F(1)]$$
$$[X \vee Y \rightarrow X] \,\&\, [X \vee Y \rightarrow Y].$$

That this is not a correct formula of the calculus of propositions may be seen by taking X and Y of opposite truth-values. Hence the formula $(Ex)F(x) \rightarrow F(y)$ is not deducible from Hilbert's postulates.

[6] *Loc. cit.*, p. 66.

[7] $A \,\&\, B$ is defined as $[\bar{A} \vee \bar{B}]$.

Similarly, the formula $F(y) \to (x)F(x)$ becomes successively:

$$(y)[F(y) \to (x)F(x)]$$
$$[F(0) \to (x)F(x)] \; \& \; [F(1) \to (x)F(x)]$$
$$[F(0) \to F(0) \; \& \; F(1)] \; \& \; [F(1) \to F(0) \; \& \; F(1)]$$
$$[X \to X \; \& \; Y] \; \& \; [Y \to X \; \& \; Y]$$

and this also fails when X and Y have opposite truth-values.

I now proceed to prove that Hilbert's postulates are independent. The proofs are given, not in the exact order in which the postulates are listed, but rather fall into two groups, depending on the nature of the proof-methods used. The first group comprises examples for postulates a), b), c), d), β), and α); and the second group for postulates e), f), and γ).

For all the independence proofs of the first group, I interpret $(E\alpha)\beta$ in a formula as meaning β, and also $(\alpha)\beta$ as meaning β; so that all apparent variables are eliminated from the formula. I then interpret all real variables as meaning a certain fixed individual, which I denote by the number 7; so that the formula now contains only propositional-variables, and function-variables with the argument 7; e. g. $F(7)$, $G(7)$, etc. I then interpret $F(7)$ as meaning a propositional-variable. Thus all formulas reduce to formulas in terms of the undefined ideas of the calculus of propositions. For example, the formula

$$(x)F(x) \vee (y)G(y) \to F(z) \vee (Ew)G(w)$$

becomes first

$$F(x) \vee G(y) \to F(z) \vee G(w)$$

and then

$$F(7) \vee G(7) \to F(7) \vee G(7)$$

and finally

$$X \vee Y \to X \vee Y.$$

The operations $X \vee Y$ and $\bar{X}$ (and hence $X \to Y$) are then defined by tables, in the usual way. And I say a formula is satisfied by the interpretation if it is such as to give the value 0 for all possible values of the propositional-variables.

Tables[8] *for Postulate* a)

$\vee$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	0

X	$\bar{X}$
0	1
1	0
2	2

$\to$	0	1	2
0	0	1	2
1	0	0	0
2	0	2	0

[8] The tables used for postulates a), c), and d) are given by Hilbert and Ackermann in another connection. *Loc. cit.*, pp. 31-33.

(Here propositional-variables can take on the values 0, 1, 2.)

Postulate a) fails when $X = 2$. For $2 \vee 2 \rightarrow 2$ becomes $0 \rightarrow 2$, which is 2.

Postulates b), c), and d) are satisfied.

Postulate e) becomes, successively:

$$F(x) \rightarrow F(y)$$
$$F(7) \rightarrow F(7)$$
$$X \rightarrow X,$$

and we see from the table for $X \rightarrow Y$ that $X \rightarrow X$ always has the value 0.

Postulate f), similarly, becomes $X \rightarrow X$.

Postulate α) is satisfied since any expression reduces to a complex involving only propositional variables and the signs $\vee$ and $\bar{\ }$; and, since the tables define class-closing operations, such a complex cannot take on *more* (though, of course, it may take on fewer) values than the propositional-variable for which it is substituted.

To see that postulate β) is satisfied, we need only observe that the table for $X \rightarrow Y$ never gives the value 0 when $X = 0$ and $Y \neq 0$.

Postulate γ) is satisfied since $A \rightarrow B(x)$ and $A \rightarrow (x)B(x)$ reduce to the same thing; and similarly for $B(x) \rightarrow A$ and $(Ex)B(x) \rightarrow A$.

Thus all the postulates except postulate a) are satisfied, so a) is independent of the other postulates.

Tables[9] *for Postulate* b)

$\vee$	0	1	2	3
0	0	0	0	0
1	0	3	0	3
2	0	0	2	2
3	0	3	2	3

X	$\bar{X}$
0	3
1	2
2	1
3	0

$\rightarrow$	0	1	2	3
0	0	3	2	3
1	0	0	2	2
2	0	3	0	3
3	0	0	0	0

(Propositional-variables take on values 0, 1, 2, 3.)

Postulate b) fails for $X = Y = 1$; since $1 \rightarrow 1 \vee 1$ becomes $1 \rightarrow 3$, which is 2.

The other postulates can be shown to hold as in the discussion for postulate a).

[9] The tables for postulate b) are isomorphic to certain tables given by Wajsberg to show the consistency of C. I. Lewis' system of "strict implication" (see Lewis and Langford, *Symbolic Logic*, p. 493, Group II), where $\bar{X}$ corresponds to $\sim p$, $X \rightarrow Y$ to $p \prec q$, and $X \vee Y$ to $\sim p \prec q$. Postulate b) yields one of the so-called "paradoxes of material implication."

Tables for Postulate c)

v	0	1	2
0	0	0	0
1	0	1	1
2	0	0	2

X	$\bar{X}$
0	1
1	2
2	0

$\rightarrow$	0	1	2
0	0	1	1
1	0	0	2
2	0	0	0

(Propositional-variables take on values 0, 1, 2.)

Postulate c) fails for $X = 2$, $Y = 1$.

Tables for Postulate d)

v	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	2	0
3	0	3	0	3

X	$\bar{X}$
0	1
1	0
2	3
3	0

$\rightarrow$	0	1	2	3
0	0	1	2	3
1	0	0	0	0
2	0	3	0	3
3	0	0	0	0

(Propositional-variables take on values 0, 1, 2, 3.)

Postulate d) fails for $X = 3$, $Y = 1$, $Z = 2$.

Tables for Postulate β)

v	0	1
0	0	0
1	0	1

X	$\bar{X}$
0	0
1	0

$\rightarrow$	0	1
0	0	0
1	0	0

(Propositional variables take on the values 0, 1.)

Postulate β) fails. For the formula,

$$(X \rightarrow Y) \rightarrow (X \vee Y)$$

gives 0 for all values of X, Y; and the formula $X \rightarrow Y$ gives 0 for all values of X, Y; but the formula $X \vee Y$ gives the value 1 for $X = Y = 1$.

The other postulates are easily seen to hold.

Tables for Postulate α)

v	0	1	2
0	0	0	1
1	0	1	1
2	0	1	1

X	$\bar{X}$
0	2
1	0
2	0

$\rightarrow$	0	1	2
0	0	1	1
1	0	0	1
2	0	0	1

(Propositional-variables take on the values 0, 1. Propositional-variables do *not* take on the value 2.)

Postulate α) fails. For it is easily seen that the formula $X \vee X \rightarrow X$ is satisfied for $X = 0, 1$; and if postulate α) held, then the formula $\bar{X} \vee \bar{X} \rightarrow \bar{X}$ (got by substituting $\bar{X}$ for X in $X \vee X \rightarrow X$) should also be satisfied. But $\bar{X} \vee \bar{X} \rightarrow \bar{X}$ is not satisfied; for if we take $X = 0$, it becomes $\bar{0} \vee \bar{0} \rightarrow \bar{0}$, which is $2 \vee 2 \rightarrow 2$, which is $1 \rightarrow 2$, which is 1. Hence α) is not satisfied.

The other postulates are readily shown to hold (remembering that the propositional-variables do not assume the value 2).

I come now to the independence-proofs of the second group. The general method used here may be described as follows. Suppose we had a set of postulates P_1, P_2 and another set Q_1, Q_2. If T represents a transformation of the undefined ideas of P_1, P_2, I designate the respective transforms of P_1 and P_2 by $T(P_1)$ and $T(P_2)$. Suppose we are able to find a transformation T_1 such that $T_1(P_1)$ is deducible from Q_1, Q_2, while $T_1(P_2)$ is independent of Q_1, Q_2; then P_2 is independent of P_1. For if P_2 were deducible from P_1, then $T_1(P_2)$ would be deducible from $T_1(P_1)$, and hence from Q_1, Q_2, contrary to hypothesis. This proof still holds, moreover, even if we suppose P_1 identical with Q_1 and P_2 identical with Q_2.

Thus, to prove the independence of postulate e), we describe a certain transformation T_e to be applied to the undefined marks of the formulas of Hilbert's system. If $T_e(A)$ is deducible from Hilbert's postulates, we say that A is satisfied; if $T_e(A)$ is independent of Hilbert's postulates, we say that A is not satisfied. (To avoid prolixity, in what follows I shall write "A is H" for "A is deducible from Hilbert's postulates.") The independence-proofs for postulates f) and γ) are similar, except that transformations T_f and T_γ are there used instead of the transformation T_e.

Independence of Postulate e)

Transformation T_e. Replace all universal-operators in the formula by the corresponding existential-operators. Thus, for example, the formula

$$(x)(y)F(x,y) \rightarrow (\exists x)(\exists y)F(x,y)$$

goes over into

$$(\exists x)(\exists y)F(x,y) \rightarrow (\exists x)(\exists y)F(x,y)$$

Postulate e) fails. For, under T_e the formula $(x)F(x) \rightarrow F(y)$ becomes $(Ex)F(x) \rightarrow F(y)$, which, by Lemma IIIa, is not H.

Postulates a), b), c), d), and f) hold, since they are invariant under T_e.

To show that postulate α) holds: Let A be any formula which is satisfied, so that $T_e(A)$ is H. Let E be an expression of the sort specified by α); it is clear that $T_e(E)$ is also an expression of the sort specified by α). Then α) states that if we substitute E in A (obtaining A_1, say) then $T_e(A_1)$ is H. But $T_e(A_1)$ can also be obtained by substituting $T_e(E)$ in $T_e(A)$. Hence since $T_e(A)$ is H, and $T_e(E)$ is an expression of the sort specified in α), it follows that $T_e(A_1)$ is H. Hence A_1 is satisfied. Hence postulate α) holds.

To show that β) holds. Suppose that A and $A \rightarrow B$ are satisfied. Then $T_e(A)$ and $T_e(A \rightarrow B)$ are H. But $T_e(A \rightarrow B)$ is the same as $T_e(A) \rightarrow T_e(B)$. Hence $T_e(B)$ is H, so B is satisfied.

To show that γ) holds. Suppose $A \to B(x)$ is satisfied. Then $T_e(A \to B(x))$, and hence $T_e(A) \to T_e(B(x))$, is H. Hence by Lemma IIa, $T_e(A) \to (Ex)T_e(B(x))$ is H. But this is the same as $T_e(A) \to T_e((x)B(x))$, so $T_e(A) \to T_e((x)B(x))$ is H; so $T_e(A \to (x)B(x))$ is H, $A \to (x)B(x)$ is satisfied, and hence the first part of postulate γ) holds. To show that the second part of γ) holds, we suppose that $B(x) \to A$ is satisfied. Then $T_e(B(x) \to A)$, and hence $T_e(B(x)) \to T_e(A)$, is H. So $(Ex)T_e(B(x)) \to T_e(A)$, and hence $T_e((Ex)B(x) \to A)$, is H. Hence $(Ex)B(x) \to A$ is satisfied.

Independence of Postulate f)

Transformation T_f. Replace all existential-operators by the corresponding universal-operators.

Postulate f) fails. For under T_f the formula $F(y) \to (Ex)F(x)$ becomes $F(y) \to (x)F(x)$, which, by Lemma IIIb, is not H.

Postulates a), b), c), d), and *e*) hold, since they are invariant under T_f.

The proof that α) and β) hold is the same as in the example for the independence of e).

The proof that γ) holds is very like the corresponding proof in the discussion for postulate e); Lemma IIb, however, is needed instead of Lemma IIa.

Independence of Postulate γ)

Transformation T_γ. Replace all the free variables in the formula by a single free variable (which is not, however, to be the same as any of the bound variables). Thus instead of $(x)F(x) \to F(y) \,\&\, F(x)$, we write $(x)F(x) \to F(y) \,\&\, F(y)$.

Postulate γ) fails. For $F(y) \to F(x)$ is satisfied, since $T_\gamma(F(y) \to F(x))$ is $F(y) \to F(y)$, and this last is H by Lemma I; but $F(y) \to (x)F(x)$ is not satisfied, since it is invariant under T_γ and, by Lemma IIIb, is not H.

Postulates a), b), c), d), e), and f) are satisfied, since they are invariant under T_γ.

The proof that postulates α) and β) are satisfied is the same as the corresponding proof in the discussion of the independence of postulate e).

This completes the proof of the independence of the nine postulates.[10]

UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

[10] It is, however, perhaps worth mentioning that the postulates could be reduced in number by introducing one of the following definitions:

$$(Ex)F(x) = \overline{[(x)\overline{F(x)}]} \quad \text{Def.}$$

or

$$(x)F(x) = \overline{[(Ex)\overline{F(x)}]} \quad \text{Def.}$$

AN UNSOLVABLE PROBLEM OF ELEMENTARY NUMBER THEORY.[1]

By Alonzo Church.

1. Introduction. There is a class of problems of elementary number theory which can be stated in the form that it is required to find an effectively calculable function f of n positive integers, such that $f(x_1, x_2, \cdots, x_n) = 2$ [2] is a necessary and sufficient condition for the truth of a certain proposition of elementary number theory involving $x_1, x_2, \cdots, x_n$ as free variables.

An example of such a problem is the problem to find a means of determining of any given positive integer n whether or not there exist positive integers x, y, z, such that $x^n + y^n = z^n$. For this may be interpreted, required to find an effectively calculable function f, such that $f(n)$ is equal to 2 if and only if there exist positive integers x, y, z, such that $x^n + y^n = z^n$. Clearly the condition that the function f be effectively calculable is an essential part of the problem, since without it the problem becomes trivial.

Another example of a problem of this class is, for instance, the problem of topology, to find a complete set of effectively calculable invariants of closed three-dimensional simplicial manifolds under homeomorphisms. This problem can be interpreted as a problem of elementary number theory in view of the fact that topological complexes are representable by matrices of incidence. In fact, as is well known, the property of a set of incidence matrices that it represent a closed three-dimensional manifold, and the property of two sets of incidence matrices that they represent homeomorphic complexes, can both be described in purely number-theoretic terms. If we enumerate, in a straightforward way, the sets of incidence matrices which represent closed three-dimensional manifolds, it will then be immediately provable that the problem under consideration (to find a complete set of effectively calculable invariants of closed three-dimensional manifolds) is equivalent to the problem, to find an effectively calculable function f of positive integers, such that $f(m, n)$ is equal to 2 if and only if the m-th set of incidence matrices and the n-th set of incidence matrices in the enumeration represent homeomorphic complexes.

Other examples will readily occur to the reader.

---

[1] Presented to the American Mathematical Society, April 19, 1935.

[2] The selection of the particular positive integer 2 instead of some other is, of course, accidental and non-essential.

The purpose of the present paper is to propose a definition of effective calculability [3] which is thought to correspond satisfactorily to the somewhat vague intuitive notion in terms of which problems of this class are often stated, and to show, by means of an example, that not every problem of this class is solvable.

2. Conversion and λ-definability. We select a particular list of symbols, consisting of the symbols { , }, (,), λ, [,], and an enumerably infinite set of symbols a, b, c, · · · to be called *variables.* And we define the word *formula* to mean any finite sequence of symbols out of this list. The terms *well-formed formula, free variable,* and *bound variable* are then defined by induction as follows. A variable $\boldsymbol{x}$ standing alone is a well-formed formula and the occurrence of $\boldsymbol{x}$ in it is an occurrence of $\boldsymbol{x}$ as a free variable in it; if the formulas $\boldsymbol{F}$ and $\boldsymbol{X}$ are well-formed, $\{\boldsymbol{F}\}(\boldsymbol{X})$ is well-formed, and an occurrence of $\boldsymbol{x}$ as a free (bound) variable in $\boldsymbol{F}$ or $\boldsymbol{X}$ is an occurrence of $\boldsymbol{x}$ as a free (bound) variable in $\{\boldsymbol{F}\}(\boldsymbol{X})$; if the formula $\boldsymbol{M}$ is well-formed and contains an occurrence of $\boldsymbol{x}$ as a free variable in $\boldsymbol{M}$, then $\lambda\boldsymbol{x}[\boldsymbol{M}]$ is well-formed, any occurrence of $\boldsymbol{x}$ in $\lambda\boldsymbol{x}[\boldsymbol{M}]$ is an occurrence of $\boldsymbol{x}$ as a bound variable in $\lambda\boldsymbol{x}[\boldsymbol{M}]$, and an occurrence of a variable $\boldsymbol{y}$, other than $\boldsymbol{x}$, as a free (bound) variable in $\boldsymbol{M}$ is an occurrence of $\boldsymbol{y}$ as a free (bound) variable in $\lambda\boldsymbol{x}[\boldsymbol{M}]$.

---

[3] As will appear, this definition of effective calculability can be stated in either of two equivalent forms, (1) that a function of positive integers shall be called effectively calculable if it is λ-definable in the sense of § 2 below, (2) that a function of positive integers shall be called effectively calculable if it is recursive in the sense of § 4 below. The notion of λ-definability is due jointly to the present author and S. C. Kleene, successive steps towards it having been taken by the present author in the *Annals of Mathematics,* vol. 34 (1933), p. 863, and by Kleene in the *American Journal of Mathematics,* vol. 57 (1935), p. 219. The notion of recursiveness in the sense of § 4 below is due jointly to Jacques Herbrand and Kurt Gödel, as is there explained. And the proof of equivalence of the two notions is due chiefly to Kleene, but also partly to the present author and to J. B. Rosser, as explained below. The proposal to identify these notions with the intuitive notion of effective calculability is first made in the present paper (but see the first footnote to § 7 below).

With the aid of the methods of Kleene (*American Journal of Mathematics,* 1935), the considerations of the present paper could, with comparatively slight modification, be carried through entirely in terms of λ-definability, without making use of the notion of recursiveness. On the other hand, since the results of the present paper were obtained, it has been shown by Kleene (see his forthcoming paper, "General recursive functions of natural numbers") that analogous results can be obtained entirely in terms of recursiveness, without making use of λ-definability. The fact, however, that two such widely different and (in the opinion of the author) equally natural definitions of effective calculability turn out to be equivalent adds to the strength of the reasons adduced below for believing that they constitute as general a characterization of this notion as is consistent with the usual intuitive understanding of it.

We shall use heavy type letters to stand for variable or undetermined formulas. And we adopt the convention that, unless otherwise stated, each heavy type letter shall represent a well-formed formula and each set of symbols standing apart which contains a heavy type letter shall represent a well-formed formula.

When writing particular well-formed formulas, we adopt the following abbreviations. A formula $\{\boldsymbol{F}\}(\boldsymbol{X})$ may be abbreviated as $\boldsymbol{F}(\boldsymbol{X})$ in any case where $\boldsymbol{F}$ is or is represented by a single symbol. A formula $\{\{\boldsymbol{F}\}(\boldsymbol{X})\}(\boldsymbol{Y})$ may be abbreviated as $\{\boldsymbol{F}\}(\boldsymbol{X}, \boldsymbol{Y})$, or, if $\boldsymbol{F}$ is or is represented by a single symbol, as $\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{Y})$. And $\{\{\{\boldsymbol{F}\}(\boldsymbol{X})\}(\boldsymbol{Y})\}(\boldsymbol{Z})$ may be abbreviated as $\{\boldsymbol{F}\}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$, or as $\boldsymbol{F}(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$, and so on. A formula $\lambda \boldsymbol{x}_1[\lambda \boldsymbol{x}_2[\cdot \cdot \cdot \lambda \boldsymbol{x}_n[\boldsymbol{M}] \cdot \cdot \cdot]]$ may be abbreviated as $\lambda \boldsymbol{x}_1 \boldsymbol{x}_2 \cdot \cdot \cdot \boldsymbol{x}_n \cdot \boldsymbol{M}$ or as $\lambda \boldsymbol{x}_1 \boldsymbol{x}_2 \cdot \cdot \cdot \boldsymbol{x}_n \boldsymbol{M}$.

We also allow ourselves at any time to introduce abbreviations of the form that a particular symbol α shall stand for a particular sequence of symbols $\boldsymbol{A}$, and indicate the introduction of such an abbreviation by the notation $\alpha \rightarrow \boldsymbol{A}$, to be read, "$\alpha$ stands for $\boldsymbol{A}$."

We introduce at once the following infinite list of abbreviations,

$$
\begin{aligned}
&1 \rightarrow \lambda ab \cdot a(b), \\
&2 \rightarrow \lambda ab \cdot a(a(b)), \\
&3 \rightarrow \lambda ab \cdot a(a(a(b))),
\end{aligned}
$$

and so on, each positive integer in Arabic notation standing for a formula of the form $\lambda ab \cdot a(a(\cdot \cdot \cdot a(b) \cdot \cdot \cdot))$.

The expression $\mathrm{S}_{\boldsymbol{N}}^{\boldsymbol{x}} \boldsymbol{M} \mid$ is used to stand for the result of substituting $\boldsymbol{N}$ for $\boldsymbol{x}$ throughout $\boldsymbol{M}$.

We consider the three following operations on well-formed formulas:

I. *To replace any part* $\lambda \boldsymbol{x}[\boldsymbol{M}]$ *of a formula by* $\lambda \boldsymbol{y}[\mathrm{S}_{\boldsymbol{y}}^{\boldsymbol{x}} \boldsymbol{M} \mid]$, *where* $\boldsymbol{y}$ *is a variable which does not occur in* $\boldsymbol{M}$.

II. *To replace any part* $\{\lambda \boldsymbol{x}[\boldsymbol{M}]\}(\boldsymbol{N})$ *of a formula by* $\mathrm{S}_{\boldsymbol{N}}^{\boldsymbol{x}} \boldsymbol{M} \mid$, *provided that the bound variables in* $\boldsymbol{M}$ *are distinct both from* $\boldsymbol{x}$ *and from the free variables in* $\boldsymbol{N}$.

III. *To replace any part* $\mathrm{S}_{\boldsymbol{N}}^{\boldsymbol{x}} \boldsymbol{M} \mid$ *(not immediately following* λ*) of a formula by* $\{\lambda \boldsymbol{x}[\boldsymbol{M}]\}(\boldsymbol{N})$, *provided that the bound variables in* $\boldsymbol{M}$ *are distinct both from* $\boldsymbol{x}$ *and from the free variables in* $\boldsymbol{N}$.

Any finite sequence of these operations is called a *conversion*, and if $\boldsymbol{B}$ is obtainable from $\boldsymbol{A}$ by a conversion we say that $\boldsymbol{A}$ is *convertible* into $\boldsymbol{B}$, or, "$\boldsymbol{A}$ conv $\boldsymbol{B}$." If $\boldsymbol{B}$ is identical with $\boldsymbol{A}$ or is obtainable from $\boldsymbol{A}$ by a single

application of one of the operations I, II, III, we say that $\boldsymbol{A}$ is *immediately convertible* into $\boldsymbol{B}$.

A conversion which contains exactly one application of Operation II, and no application of Operation III, is called a *reduction.*

A formula is said to be *in normal form* if it is well-formed and contains no part of the form $\{\lambda \boldsymbol{x}[\boldsymbol{M}]\}(\boldsymbol{N})$. And $\boldsymbol{B}$ is said to be a *normal form of* $\boldsymbol{A}$ if $\boldsymbol{B}$ is in normal form and $\boldsymbol{A}$ conv $\boldsymbol{B}$.

The originally given order $a, b, c, \cdots$ of the variables is called their *natural order.* And a formula is said to be *in principal normal form* if it is in normal form, and no variable occurs in it both as a free variable and as a bound variable, and the variables which occur in it immediately following the symbol λ are, when taken in the order in which they occur in the formula, in natural order without repetitions, beginning with a and omitting only such variables as occur in the formula as free variables.[4] The formula $\boldsymbol{B}$ is said to be the *principal normal form of* $\boldsymbol{A}$ if $\boldsymbol{B}$ is in principal normal form and $\boldsymbol{A}$ conv $\boldsymbol{B}$.

Of the three following theorems, proof of the first is immediate, and the second and third have been proved by the present author and J. B. Rosser:[5]

THEOREM I. *If a formula is in normal form, no reduction of it is possible.*

THEOREM II. *If a formula has a normal form, this normal form is unique to within applications of Operation* I, *and any sequence of reductions of the formula must (if continued) terminate in the normal form.*

THEOREM III. *If a formula has a normal form, every well-formed part of it has a normal form.*

We shall call a function a *function of positive integers* if the range of each independent variable is the class of positive integers and the range of the dependent variable is contained in the class of positive integers. And when it is desired to indicate the number of independent variables we shall speak of a function of one positive integer, a function of two positive integers, and so on. Thus if F is a function of n positive integers, and $a_1, a_2, \cdots, a_n$ are positive integers, then $F(a_1, a_2, \cdots, a_n)$ must be a positive integer.

---

[4] For example, the formulas $\lambda ab \cdot b(a)$ and $\lambda a \cdot a(\lambda c \cdot b(c))$ are in principal normal form, and $\lambda ac \cdot c(a)$, and $\lambda bc \cdot c(b)$, and $\lambda a \cdot a(\lambda a \cdot b(a))$ are in normal form but not in principal normal form. Use of the principal normal form was suggested by S. C. Kleene as a means of avoiding the ambiguity of determination of the normal form of a formula, which is troublesome in certain connections.

Observe that the formulas $1, 2, 3, \cdots$ are all in principal normal form.

[5] Alonzo Church and J. B. Rosser, "Some properties of conversion," forthcoming (abstract in *Bulletin of the American Mathematical Society*, vol. 41, p. 332).

A function F of one positive integer is said to be *λ-definable* if it is possible to find a formula $\boldsymbol{F}$ such that, if $F(m) = r$ and $\boldsymbol{m}$ and $\boldsymbol{r}$ are the formulas for which the positive integers m and r (written in Arabic notation) stand according to our abbreviations introduced above, then $\{\boldsymbol{F}\}(\boldsymbol{m})$ conv $\boldsymbol{r}$.

Similarly, a function F of two positive integers is said to be λ-definable if it is possible to find a formula $\boldsymbol{F}$ such that, whenever $F(m, n) = r$, the formula $\{\boldsymbol{F}\}(\boldsymbol{m}, \boldsymbol{n})$ is convertible into $\boldsymbol{r}$ (m, n, r being positive integers and $\boldsymbol{m}$, $\boldsymbol{n}$, $\boldsymbol{r}$ the corresponding formulas). And so on for functions of three or more positive integers.[6]

It is clear that, in the case of any λ-definable function of positive integers, the process of reduction of formulas to normal form provides an algorithm for the effective calculation of particular values of the function.

3. The Gödel representation of a formula. Adapting to the formal notation just described a device which is due to Gödel,[7] we associate with every formula a positive integer to represent it, as follows. To each of the symbols {, (, [we let correspond the number 11, to each of the symbols },),] the number 13, to the symbol λ the number 1, and to the variables a, b, c, $\cdots$ the prime numbers 17, 19, 23, $\cdots$ respectively. And with a formula which is composed of the n symbols $\tau_1, \tau_2, \cdots, \tau_n$ in order we associate the number $2^{t_1}3^{t_2} \cdots p_n{}^{t_n}$, where t_i is the number corresponding to the symbol τ_i, and where p_n stands for the n-th prime number.

This number $2^{t_1}3^{t_2} \cdots p_n{}^{t_n}$ will be called the *Gödel representation* of the formula $\tau_1\tau_2 \cdots \tau_n$.

Two distinct formulas may sometimes have the same Gödel representation, because the numbers 11 and 13 each correspond to three different symbols, but it is readily proved that *no two distinct well-formed formulas can have the same Gödel representation.* It is clear, moreover, that there is an effective method by which, given any formula, its Gödel representation can be calculated; and likewise that there is an effective method by which, given any positive integer, it is possible to determine whether it is the Gödel representation of a well-formed formula and, if it is, to obtain that formula.

In this connection the Gödel representation plays a rôle similar to that

---

[6] Cf. S. C. Kleene, "A theory of positive integers in formal logic," *American Journal of Mathematics,* vol. 57 (1935), pp. 153-173 and 219-244, where the λ-definability of a number of familiar functions of positive integers, and of a number of important general classes of functions, is established. Kleene uses the term *definable,* or *formally definable,* in the sense in which we are here using *λ-definable.*

[7] Kurt Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I," *Monatshefte für Mathematik und Physik,* vol. 38 (1931), pp. 173-198.

of the matrix of incidence in combinatorial topology (cf. § 1 above). For there is, in the theory of well-formed formulas, an important class of problems, each of which is equivalent to a problem of elementary number theory obtainable by means of the Gödel representation.[8]

4. Recursive functions. We define a class of expressions, which we shall call *elementary expressions,* and which involve, besides parentheses and commas, the symbols 1, S, an infinite set of numerical variables $x, y, z, \cdots$, and, for each positive integer n, an infinite set $f_n, g_n, h_n, \cdots$ of functional variables with subscript n. This definition is by induction as follows. The symbol 1 or any numerical variable, standing alone, is an elementary expression. If A is an elementary expression, then $S(A)$ is an elementary expression. If $A_1, A_2, \cdots, A_n$ are elementary expressions and f_n is any functional variable with subscript n, then $f_n(A_1, A_2, \cdots, A_n)$ is an elementary expression.

The particular elementary expressions 1, $S(1)$, $S(S(1)), \cdots$ are called *numerals.* And the positive integers $1, 2, 3, \cdots$ are said to correspond to the numerals 1, $S(1)$, $S(S(1)), \cdots$.

An expression of the form $A = B$, where A and B are elementary expressions, is called an *elementary equation.*

The *derived equations* of a set E of elementary equations are defined by induction as follows. The equations of E themselves are derived equations. If $A = B$ is a derived equation containing a numerical variable x, then the result of substituting a particular numeral for all the occurrences of x in $A = B$ is a derived equation. If $A = B$ is a derived equation containing an elementary expression C (as part of either A or B), and if either $C = D$ or $D = C$ is a derived equation, then the result of substituting D for a particular occurrence of C in $A = B$ is a derived equation.

Suppose that no derived equation of a certain finite set E of elementary equations has the form $k = l$ where k and l are different numerals, that the functional variables which occur in E are $f_{n_1}{}^1, f_{n_2}{}^2, \cdots, f_{n_r}{}^r$ with subscripts $n_1, n_2, \cdots, n_r$ respectively, and that, for every value of i from 1 to r inclusive, and for every set of numerals $k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i$, there exists a unique numeral k^i such that $f_{n_i}{}^i(k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i) = k^i$ is a derived equation of E. And let $F^1, F^2, \cdots, F^r$ be the functions of positive integers defined by the con-

[8] This is merely a special case of the now familiar remark that, in view of the Gödel representation and the ideas associated with it, symbolic logic in general can be regarded, mathematically, as a branch of elementary number theory. This remark is essentially due to Hilbert (cf. for example, *Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg,* 1904, p. 185; also Paul Bernays in *Die Naturwissenschaften,* vol. 10 (1922), pp. 97 and 98) but is most clearly formulated in terms of the Gödel representation.

dition that, in all cases, $F^i(m_1{}^i, m_2{}^i, \cdots, m_{n_i}{}^i)$ shall be equal to m^i, where $m_1{}^i, m_2{}^i, \cdots, m_{n_i}{}^i$, and m^i are the positive integers which correspond to the numerals $k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i$, and k^i respectively. Then the set of equations E is said to *define*, or to be a set of *recursion equations* for, any one of the functions F^i, and the functional variable $f_{n_i}{}^i$ is said to *denote* the function F^i.

A function of positive integers for which a set of recursion equations can be given is said to be *recursive*.[9]

It is clear that for any recursive function of positive integers there exists an algorithm using which any required particular value of the function can be effectively calculated. For the derived equations of the set of recursion equations E are effectively enumerable, and the algorithm for the calculation of particular values of a function F^i, denoted by a functional variable $f_{n_i}{}^i$, consists in carrying out the enumeration of the derived equations of E until the required particular equation of the form $f_{n_i}{}^i(k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i) = k^i$ is found.[10]

We call an infinite sequence of positive integers recursive if the function F such that $F(n)$ is the n-th term of the sequence is recursive.

We call a propositional function of positive integers recursive if the function whose value is 2 or 1, according to whether the propositional function is true or false, is recursive. By a recursive property of positive integers we shall mean a recursive propositional function of one positive integer, and by a recursive relation between positive integers we shall mean a recursive propositional function of two or more positive integers.

---

[9] This definition is closely related to, and was suggested by, a definition of recursive functions which was proposed by Kurt Gödel, in lectures at Princeton, N. J., 1934, and credited by him in part to an unpublished suggestion of Jacques Herbrand. The principal features in which the present definition of recursiveness differs from Gödel's are due to S. C. Kleene.

In a forthcoming paper by Kleene to be entitled, "General recursive functions of natural numbers," (abstract in *Bulletin of the American Mathematical Society*, vol. 41), several definitions of recursiveness will be discussed and equivalences among them obtained. In particular, it follows readily from Kleene's results in that paper that every function recursive in the present sense is also recursive in the sense of Gödel (1934) and conversely.

[10] The reader may object that this algorithm cannot be held to provide an effective calculation of the required particular value of F^i unless the proof is constructive that the required equation $f_{n_i}{}^i(k_1{}^i, k_2{}^i, \cdots, k_{n_i}{}^i) = k^i$ will ultimately be found. But if so this merely means that he should take the existential quantifier which appears in our definition of a set of recursion equations in a constructive sense. What the criterion of constructiveness shall be is left to the reader.

The same remark applies in connection with the existence of an algorithm for calculating the values of a λ-definable function of positive integers.

A function F, for which the range of the dependent variable is contained in the class of positive integers and the range of the independent variable, or of each independent variable, is a subset (not necessarily the whole) of the class of positive integers, will be called *potentially recursive,* if it is possible to find a recursive function F' of positive integers (for which the range of the independent variable, or of each independent variable, is the whole of the class of positive integers), such that the value of F' agrees with the value of F in all cases where the latter is defined.

By an *operation on* well-formed formulas we shall mean a function for which the range of the dependent variable is contained in the class of well-formed formulas and the range of the independent variable, or of each independent variable, is the whole class of well-formed formulas. And we call such an operation recursive if the corresponding function obtained by replacing all formulas by their Gödel representations is potentially recursive.

Similarly any function for which the range of the dependent variable is contained either in the class of positive integers or in the class of well-formed formulas, and for which the range of each independent variable is identical either with the class of positive integers or with the class of well-formed formulas (allowing the case that some of the ranges are identical with one class and some with the other), will be said to be recursive if the corresponding function obtained by replacing all formulas by their Gödel representations is potentially recursive. We call an infinite sequence of well-formed formulas recursive if the corresponding infinite sequence of Gödel representations is recursive. And we call a property of, or relation between, well-formed formulas recursive if the corresponding property of, or relation between, their Gödel representations is potentially recursive. A set of well-formed formulas is said to be recursively enumerable if there exists a recursive infinite sequence which consists entirely of formulas of the set and contains every formula of the set at least once.[11]

In terms of the notion of recursiveness we may also define a *proposition of elementary number theory,* by induction as follows. If ϕ is a recursive propositional function of n positive integers (defined by giving a particular set of recursion equations for the corresponding function whose values are 2 and 1) and if $x_1, x_2, \cdots, x_n$ are variables which take on positive integers as values, then $\phi(x_1, x_2, \cdots, x_n)$ is a proposition of elementary number theory. If P is a proposition of elementary number theory involving x as a free

[11] It can be shown, in view of Theorem V below, that, if an infinite set of formulas is recursively enumerable in this sense, it is also recursively enumerable in the sense that there exists a recursive infinite sequence which consists entirely of formulas of the set and contains every formula of the set exactly once.

variable, then the result of substituting a particular positive integer for all occurrences of x as a free variable in P is a proposition of elementary number theory, and $(x)P$ and $(\exists x)P$ are propositions of elementary number theory, where (x) and $(\exists x)$ are respectively the universal and existential quantifiers of x over the class of positive integers.

It is then readily seen that the negation of a proposition of elementary number theory or the logical product or the logical sum of two propositions of elementary number theory is equivalent, in a simple way, to another proposition of elementary number theory.

5. Recursiveness of the Kleene $\mathfrak{p}$-function. We prove two theorems which establish the recursiveness of certain functions which are definable in words by means of the phrase, "The least positive integer such that," or, "The n-th positive integer such that."

THEOREM IV. *If F is a recursive function of two positive integers, and if for every positive integer x there exists a positive integer y such that $F(x, y) > 1$, then the function F^*, such that, for every positive integer x, $F^*(x)$ is equal to the least positive integer y for which $F(x, y) > 1$, is recursive.*

For a set of recursion equations for F^* consists of the recursion equations for F together with the equations,

$$
\begin{array}{ll}
i_2(1, 2) = 2, & g_2(x, 1) = i_2(f_2(x, 1), 2), \\
i_2(S(x), 2) = 1, & g_2(x, S(y)) = i_2(f_2(x, S(y)), g_2(x, y)), \\
i_2(x, 1) = 3, & h_2(S(x), y) = x, \\
i_2(x, S(S(y))) = 3, & h_2(g_2(x, y), x) = j_2(g_2(x, y), y), \\
j_2(1, y) = y, & f_1(x) = h_2(1, x), \\
j_2(S(x), y) = x, &
\end{array}
$$

where the functional variables f_2 and f_1 denote the functions F and F^* respectively, and 2 and 3 are abbreviations for $S(1)$ and $S(S(1))$ respectively.[12]

THEOREM V. *If F is a recursive function of one positive integer, and if there exist an infinite number of positive integers x for which $F(x) > 1$, then the function F^0, such that, for every positive integer n, $F^0(n)$ is equal to the n-th positive integer x (in order of increasing magnitude) for which $F(x) > 1$, is recursive.*

---

[12] Since this result was obtained, it has been pointed out to the author by S. C. Kleene that it can be proved more simply by using the methods of the latter in *American Journal of Mathematics*, vol. 57 (1935), p. 231 *et seq.* His proof will be given in his forthcoming paper already referred to.

For a set of recursion equations for F^0 consists of the recursion equations for F together with the equations,

$$\begin{aligned} g_2(1, y) &= g_2(f_1(S(y)), S(y)), \\ g_2(S(x), y) &= y, \\ g_1(1) &= k, \\ g_1(S(y)) &= g_2(1, g_1(y)), \end{aligned}$$

where the functional variables g_1 and f_1 denote the functions F^0 and F respectively, and where k is the numeral to which corresponds the least positive integer x for which $F(x) > 1$.[13]

6. Recursiveness of certain functions of formulas. We list now a number of theorems which will be proved in detail in a forthcoming paper by S. C. Kleene[14] or follow immediately from considerations there given. We omit proofs here, except for brief indications in some instances.

Our statement of the theorems and our notation differ from Kleene's in that we employ the set of positive integers $(1, 2, 3, \cdots)$ in the rôle in which he employs the set of natural numbers $(0, 1, 2, \cdots)$. This difference is, of course, unessential. We have selected what is, from some points of view, the less natural alternative, in order to preserve the convenience and naturalness of the identification of the formula $\lambda ab \cdot a(b)$ with 1 rather than with 0.

THEOREM VI. *The property of a positive integer, that there exists a well-formed formula of which it is the Gödel representation is recursive.*

THEOREM VII. *The set of well-formed formulas is recursively enumerable.*

This follows from Theorems V and VI.

THEOREM VIII. *The function of two variables, whose value, when taken of the well-formed formulas* $\mathbf{F}$ *and* $\mathbf{X}$, *is the formula* $\{\mathbf{F}\}(\mathbf{X})$, *is recursive.*

THEOREM IX. *The function, whose value for each of the positive integers* $1, 2, 3, \cdots$ *is the corresponding formula* $1, 2, 3, \cdots$, *is recursive.*

THEOREM X. *A function, whose value for each of the formulas* $1, 2, 3, \cdots$ *is the corresponding positive integer, and whose value for other well-formed formulas is a fixed positive integer, is recursive. Likewise the function, whose value for each of the formulas* $1, 2, 3, \cdots$ *is the corresponding positive integer*

[13] This proof is due to Kleene.

[14] S. C. Kleene, "λ-definability and recursiveness," forthcoming (abstract in *Bulletin of the American Mathematical Society*, vol. 41). In connection with many of the theorems listed, see also Kurt Gödel, *Monatshefte für Mathematik und Physik*, vol. 38 (1931), p. 181 *et seq.*, observing that every function which is recursive in the sense in which the word is there used by Gödel is also recursive in the present more general sense.

plus one, and whose value for other well-formed formulas is the positive integer 1, *is recursive.*

THEOREM XI. *The relation of immediate convertibility, between well-formed formulas, is recursive.*

THEOREM XII. *It is possible to associate simultaneously with every well-formed formula an enumeration of the formulas obtainable from it by conversion, in such a way that the function of two variables, whose value, when taken of a well-formed formula A and a positive integer n, is the n-th formula in the enumeration of the formulas obtainable from A by conversion, is recursive.*

THEOREM XIII. *The property of a well-formed formula, that it is in principal normal form, is recursive.*

THEOREM XIV. *The set of well-formed formulas which are in principal normal form is recursively enumerable.*

This follows from Theorems V, VII, XIII.

THEOREM XV. *The set of well-formed formulas which have a normal form is recursively enumerable.*[15]

For by Theorems XII and XIV this set can be arranged in an infinite square array which is recursively defined (i. e. defined by a recursive function of two variables). And the familiar process by which this square array is reduced to a single infinite sequence is recursive (i. e. can be expressed by means of recursive functions).

THEOREM XVI. *Every recursive function of positive integers is λ-definable.*[16]

THEOREM XVII. *Every λ-definable function of positive integers is recursive.*[17]

For functions of one positive integer this follows from Theorems IX, VIII, XII, XIII, IV, X. For functions of more than one positive integer

---

[15] This theorem was first proposed by the present author, with the outline of proof here indicated. Details of its proof are due to Kleene and will be given by him in his forthcoming paper, "λ-definability and recursiveness."

[16] This theorem can be proved as a straightforward application of the methods introduced by Kleene in the *American Journal of Mathematics* (*loc. cit.*). In the form here given it was first obtained by Kleene. The related result had previously been obtained by J. B. Rosser that, if we modify the definition of *well-formed* by omitting the requirement that M contain x as a free variable in order that $\lambda x[M]$ be well-formed, then every recursive function of positive integers is λ-definable in the resulting modified sense.

[17] This result was obtained independently by the present author and S. C. Kleene at about the same time.

it follows by the same method, using a generalization of Theorem IV to functions of more than two positive integers.

7. The notion of effective calculability. We now define the notion, already discussed, of an *effectively calculable* function of positive integers by identifying it with the notion of a recursive function of positive integers [18] (or of a λ-definable function of positive integers). This definition is thought to be justified by the considerations which follow, so far as positive justification can ever be obtained for the selection of a formal definition to correspond to an intuitive notion.

It has already been pointed out that, for every function of positive integers which is effectively calculable in the sense just defined, there exists an algorithm for the calculation of its values.

Conversely it is true, under the same definition of effective calculability, that every function, an algorithm for the calculation of the values of which exists, is effectively calculable. For example, in the case of a function F of one positive integer, an algorithm consists in a method by which, given any positive integer n, a sequence of expressions (in some notation) $E_{n1}, E_{n2}, \cdots, E_{nr_n}$, can be obtained; where E_{n1} is effectively calculable when n is given; where E_{ni} is effectively calculable when n and the expressions E_{nj}, $j < i$, are given; and where, when n and all the expressions E_{ni} up to and including E_{nr_n} are given, the fact that the algorithm has terminated becomes effectively known and the value of $F(n)$ is effectively calculable. Suppose that we set up a system of Gödel representations for the notation employed in the expressions E_{ni}, and that we then further adopt the method of Gödel of representing a finite sequence of expressions $E_{n1}, E_{n2}, \cdot \cdot \cdot, E_{ni}$ by the single positive integer $2^{e_{n1}}3^{e_{n2}} \cdot \cdot \cdot p_i^{e_{ni}}$ where $e_{n1}, e_{n2}, \cdot \cdot \cdot, e_{ni}$ are respectively the Gödel representations of $E_{n1}, E_{n2}, \cdot \cdot \cdot, E_{ni}$ (in particular representing a vacuous sequence of expressions by the positive integer 1). Then we may define a function G of two positive integers such that, if x represents the finite sequence $E_{n1}, E_{n2}, \cdot \cdot \cdot, E_{nk}$, then $G(n, x)$ is equal to the Gödel representation of E_{ni}, where $i = k + 1$, or is equal to 10 if $k = r_n$ (that is if the algorithm has terminated with E_{nk}), and in any other case $G(n, x)$ is equal to 1. And we may define a function H of two positive integers, such that the value of $H(n, x)$ is the same as that of $G(n, x)$, except in the case that $G(n, x) = 10$, in which case $H(n, x) = F(n)$. If the interpretation is allowed that the

[18] The question of the relationship between effective calculability and recursiveness (which it is here proposed to answer by identifying the two notions) was raised by Gödel in conversation with the author. The corresponding question of the relationship between effective calculability and λ-definability had previously been proposed by the author independently.

requirement of effective calculability which appears in our description of an algorithm means the effective calculability of the functions G and H,[19] and if we take the effective calculability of G and H to mean recursiveness (λ-definability), then the recursiveness (λ-definability) of F follows by a straightforward argument.

Suppose that we are dealing with some particular system of symbolic logic, which contains a symbol, $=$, for equality of positive integers, a symbol $\{\ \}(\ )$ for the application of a function of one positive integer to its argument, and expressions $1, 2, 3, \cdots$ to stand for the positive integers. The theorems of the system consist of a finite, or enumerably infinite, list of expressions, the *formal axioms,* together with all the expressions obtainable from them by a finite succession of applications of operations chosen out of a given finite, or enumerably infinite, list of operations, the *rules of procedure.* If the system is to serve at all the purposes for which a system of symbolic logic is usually intended, it is necessary that each rule of procedure be an effectively calculable operation, that the complete set of rules of procedure (if infinite) be effectively enumerable, that the complete set of formal axioms (if infinite) be effectively enumerable, and that the relation between a positive integer and the expression which stands for it be effectively determinable. Suppose that we interpret this to mean that, in terms of a system of Gödel representations for the expressions of the logic, each rule of procedure must be a recursive operation,[20] the complete set of rules of procedure must be recursively enumerable (in the sense that there exists a recursive function Φ such that $\Phi(n, x)$ is the representation of the result of applying the n-th rule of procedure to the ordered finite set of formulas represented by x), the complete set of formal axioms must be recursively enumerable, and the relation between a positive integer and the expression which stands for it must be recursive.[21] And let us call a function F of one positive integer[22] *calculable within* the logic if there exists an expression f in the logic such that $\{f\}(\mu) = \nu$ is a theorem when and only when $F(m) = n$ is true, μ and ν being the expressions which stand for the positive integers m and n. Then, since the

[19] If this interpretation or some similar one is not allowed, it is difficult to see how the notion of an algorithm can be given any exact meaning at all.

[20] As a matter of fact, in known systems of symbolic logic, e. g. in that of *Principia Mathematica,* the stronger statement holds, that the relation of *immediate consequence* (*unmittelbare Folge*) is recursive. Cf. Gödel, *loc. cit.,* p. 185. In any case where the relation of immediate consequence is recursive it is possible to find a set of rules of procedure, equivalent to the original ones, such that each rule is a (one-valued) recursive operation, and the complete set of rules is recursively enumerable.

[21] The author is here indebted to Gödel, who, in his 1934 lectures already referred to, proposed substantially these conditions, but in terms of the more restricted notion

complete set of theorems of the logic is recursively enumerable, it follows by Theorem IV above that every function of one positive integer which is calculable within the logic is also effectively calculable (in the sense of our definition).

Thus it is shown that no more general definition of effective calculability than that proposed above can be obtained by either of two methods which naturally suggest themselves (1) by defining a function to be effectively calculable if there exists an algorithm for the calculation of its values (2) by defining a function F (of one positive integer) to be effectively calculable if, for every positive integer m, there exists a positive integer n such that $F(m) = n$ is a provable theorem.

8. Invariants of conversion. The problem naturally suggests itself to find invariants of that transformation of formulas which we have called conversion. The only effectively calculable invariants at present known are the immediately obvious ones (e. g. the set of free variables contained in a formula). Others of importance very probably exist. But we shall prove (in Theorem XIX) that, under the definition of effective calculability proposed in § 7, *no complete set of effectively calculable invariants of conversion exists* (cf. § 1).

The results of Kleene (*American Journal of Mathematics*, 1935) make it clear that, if the problem of finding a complete set of effectively calculable invariants of conversion were solved, most of the familiar unsolved problems of elementary number theory would thereby also be solved. And from Theorem XVI above it follows further that to find a complete set of effectively calculable invariants of conversion would imply the solution of the Entscheidungsproblem for any system of symbolic logic whatever (subject to the very general restrictions of § 7). In the light of this it is hardly surprising that the problem to find such a set of invariants should be unsolvable.

It is to be remembered, however, that, if we consider only the statement of the problem (and ignore things which can be proved about it by more or less lengthy arguments), it appears to be a problem of the same class as the problems of number theory and topology to which it was compared in § 1, having no striking characteristic by which it can be distinguished from them. The temptation is strong to reason by analogy that other important problems of this class may also be unsolvable.

---

of recursiveness which he had employed in 1931, and using the condition that the relation of immediate consequence be recursive instead of the present conditions on the rules of procedure.

[22] We confine ourselves for convenience to the case of functions of one positive integer. The extension to functions of several positive integers is immediate.

LEMMA. *The problem, to find a recursive function of two formulas* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *whose value is* 2 *or* 1 *according as* $\boldsymbol{A}$ *conv* $\boldsymbol{B}$ *or not, is equivalent to the problem, to find a recursive function of one formula* $\boldsymbol{C}$ *whose value is* 2 *or* 1 *according as* $\boldsymbol{C}$ *has a normal form or not.*[23]

For, by Theorem X, the formula $\boldsymbol{a}$ (the formula $\boldsymbol{b}$), which stands for the positive integer which is the Gödel representation of the formula $\boldsymbol{A}$ (the formula $\boldsymbol{B}$), can be expressed as a recursive function of the formula $\boldsymbol{A}$ (the formula $\boldsymbol{B}$). Moreover, by Theorems VI and XII, there exists a recursive function F of two positive integers such that, if m is the Gödel representation of a well-formed formula $\boldsymbol{M}$, then $F(m, n)$ is the Gödel representation of the n-th formula in an enumeration of the formulas obtainable from $\boldsymbol{M}$ by conversion. And, by Theorem XVI, F is λ-definable, by a formula $\mathfrak{f}$. If we define,

$$Z_1 \rightarrow \mathfrak{Q}(\lambda x \cdot x(I), I),$$
$$Z_2 \rightarrow \mathfrak{Q}(\lambda xy \cdot S(x) - y, I),$$

where $\mathfrak{Q}$ is the formula defined by Kleene (*American Journal of Mathematics*, vol. 57 (1935), p. 226), then Z_1 and Z_2 λ-define the functions of one positive integer whose values, for a positive integer n, are the n-th terms respectively of the infinite sequences $1, 1, 2, 1, 2, 3, \cdots$ and $1, 2, 1, 3, 2, 1, \cdots$. By Theorem VIII the formula,

$$\{\lambda xy \cdot \mathfrak{p}(\lambda n \cdot \delta(\mathfrak{f}(x, Z_1(n)), \mathfrak{f}(y, Z_2(n))), 1)\}(\boldsymbol{a}, \boldsymbol{b}),$$

where $\mathfrak{p}$ and δ are defined as by Kleene (*loc. cit.*, p. 173 and p. 231), is a recursive function of $\boldsymbol{A}$ and $\boldsymbol{B}$, and this formula has a normal form if and only if $\boldsymbol{A}$ conv $\boldsymbol{B}$.

Again, by Theorem X, the formula $\boldsymbol{c}$, which stands for the positive integer which is the Gödel representation of the formula $\boldsymbol{C}$, can be expressed as a recursive function of the formula $\boldsymbol{C}$. By Theorems VI and XIII, there exists a recursive function G of one positive integer such that $G(m) = 2$ if m is the Gödel representation of a formula in principal normal form, and $G(m) = 1$ in any other case. And, by Theorem XVI, G is λ-definable, by a formula $\mathfrak{g}$. By Theorem VIII the formula,

$$\{\lambda x \cdot \mathfrak{p}(\lambda n \cdot \mathfrak{g}(\mathfrak{f}(x, n), 1, 1))\}(\boldsymbol{c})$$

[23] These two problems, in the forms, (1) to find an effective method of determining of any two formulas A and B whether A conv B, (2) to find an effective method of determining of any formula C whether it has a normal form, were both proposed by Kleene to the author, in the course of a discussion of the properties of the $\mathfrak{p}$-function, about 1932. Some attempts towards solution of (1) by means of numerical invariants were actually made by Kleene at about that time.

where $\mathfrak{f}$ is the formula $\mathfrak{f}$ used in the preceding paragraph, is a recursive function of $\boldsymbol{C}$, and this formula is convertible into the formula 1 if and only if $\boldsymbol{C}$ has a normal form.

Thus we have proved that a formula $\boldsymbol{C}$ can be found as a recursive function of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$, such that $\boldsymbol{C}$ has a normal form if and only if $\boldsymbol{A}$ conv $\boldsymbol{B}$; and that a formula $\boldsymbol{A}$ can be found as a recursive function of a formula $\boldsymbol{C}$, such that $\boldsymbol{A}$ conv 1 if and only if $\boldsymbol{C}$ has a normal form. From this the lemma follows.

THEOREM XVIII. *There is no recursive function of a formula* $\boldsymbol{C}$*, whose value is* 2 *or* 1 *according as* $\boldsymbol{C}$ *has a normal form or not.*

That is, the property of a well-formed formula, that it has a normal form, is not recursive.

For assume the contrary.

Then there exists a recursive function H of one positive integer such that $H(m) = 2$ if m is the Gödel representation of a formula which has a normal form, and $H(m) = 1$ in any other case. And, by Theorem XVI, H is λ-definable by a formula $\mathfrak{h}$.

By Theorem XV, there exists an enumeration of the well-formed formulas which have a normal form, and a recursive function A of one positive integer such that $A(n)$ is the Gödel representation of the n-th formula in this enumeration. And, by Theorem XVI, A is λ-definable, by a formula $\mathfrak{a}$.

By Theorems VI and VIII, there exists a recursive function B of two positive integers such that, if m and n are Gödel representations of well-formed formulas $\boldsymbol{M}$ and $\boldsymbol{N}$, then $B(m, n)$ is the Gödel representation of $\{\boldsymbol{M}\}(\boldsymbol{N})$. And, by Theorem XVI, B is λ-definable, by a formula $\mathfrak{b}$.

By Theorems VI and X, there exists a recursive function C of one positive integer such that, if m is the Gödel representation of one of the formulas $1, 2, 3, \cdots$, then $C(m)$ is the corresponding positive integer plus one, and in any other case $C(m) = 1$. And, by Theorem XVI, C is λ-definable, by a formula $\mathfrak{c}$.

By Theorem IX there exists a recursive function Z^{-1} of one positive integer, whose value for each of the positive integers $1, 2, 3, \cdots$ is the Gödel representation of the corresponding formula $1, 2, 3, \cdots$. And, by Theorem XVI, Z^{-1} is λ-definable, by a formula $\mathfrak{z}$.

Let $\mathfrak{f}$ and $\mathfrak{g}$ be the formulas $\mathfrak{f}$ and $\mathfrak{g}$ used in the proof of the Lemma. By Kleene 15III Cor. (*loc. cit.*, p. 220), a formula $\mathfrak{d}$ can be found such that,

$$\mathfrak{d}(1) \text{ conv } \lambda x \cdot x(1)$$
$$\mathfrak{d}(2) \text{ conv } \lambda u \cdot \mathfrak{c}(\mathfrak{f}(u, \mathfrak{p}(\lambda m \cdot \mathfrak{g}(\mathfrak{f}(u, m)), 1))).$$

We define,

$$\mathfrak{e} \rightarrow \lambda n \cdot \mathfrak{d}(\mathfrak{h}(\mathfrak{b}(\mathfrak{a}(n), \mathfrak{z}(n))), \mathfrak{b}(\mathfrak{a}(n), \mathfrak{z}(n))).$$

Then if $\mathbf{n}$ is one of the formulas $1, 2, 3, \cdots$, $\mathfrak{e}(\mathbf{n})$ is convertible into one of the formulas $1, 2, 3, \cdots$ in accordance with the following rules: (1) if $\mathfrak{b}(\mathfrak{a}(\mathbf{n}), \mathfrak{z}(\mathbf{n}))$ conv a formula which stands for the Gödel representation of a formula which has no normal form, $\mathfrak{e}(\mathbf{n})$ conv 1, (2) if $\mathfrak{b}(\mathfrak{a}(\mathbf{n}), \mathfrak{z}(\mathbf{n}))$ conv a formula which stands for the Gödel representation of a formula which has a principal normal form which is not one of the formulas $1, 2, 3, \cdots$, $\mathfrak{e}(\mathbf{n})$ conv 1, (3) if $\mathfrak{b}(\mathfrak{a}(\mathbf{n}), \mathfrak{z}(\mathbf{n}))$ conv a formula which stands for the Gödel representation of a formula which has a principal normal form which is one of the formulas $1, 2, 3, \cdots$, $\mathfrak{e}(\mathbf{n})$ conv the next following formula in the list $1, 2, 3, \cdots$.

By Theorem III, since $\mathfrak{e}(1)$ has a normal form, the formula $\mathfrak{e}$ has a normal form. Let $\mathfrak{E}$ be the formula which stands for the Gödel representation of $\mathfrak{e}$. Then, if $\mathbf{n}$ is any one of the formulas $1, 2, 3, \cdots$, $\mathfrak{E}$ is not convertible into the formula $\mathfrak{a}(\mathbf{n})$, because $\mathfrak{b}(\mathfrak{E}, \mathfrak{z}(\mathbf{n}))$ is, by the definition of $\mathfrak{b}$, convertible into the formula which stands for the Gödel representation of $\mathfrak{e}(\mathbf{n})$, while $\mathfrak{b}(\mathfrak{a}(\mathbf{n}), \mathfrak{z}(\mathbf{n}))$ is, by the preceding paragraph, convertible into the formula stands for the Gödel representation of a formula definitely not convertible into $\mathfrak{e}(\mathbf{n})$ (Theorem II). But, by our definition of $\mathfrak{a}$, it must be true of one of the formulas $\mathbf{n}$ in the list $1, 2, 3, \cdots$ that $\mathfrak{a}(\mathbf{n})$ conv $\mathfrak{E}$.

Thus, since our assumption to the contrary has led to a contradiction, the theorem must be true.

In order to present the essential ideas without any attempt at exact statement, the preceding proof may be outlined as follows. We are to deduce a contradiction from the assumption that it is effectively determinable of every well-formed formula whether or not it has a normal form. If this assumption holds, it is effectively determinable of every well-formed formula whether or not it is convertible into one of the formulas $1, 2, 3, \cdots$; for, given a well-formed formula $\boldsymbol{R}$, we can first determine whether or not it has a normal form, and if it has we can obtain the principal normal form by enumerating the formulas into which $\boldsymbol{R}$ is convertible (Theorem XII) and picking out the first formula in principal normal form which occurs in the enumeration, and we can then determine whether the principal normal form is one of the formulas $1, 2, 3, \cdots$. Let $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \cdots$ be an effective enumeration of the well-formed formulas which have a normal form (Theorem XV). Let E be a function of one positive integer, defined by the rule that, where $\boldsymbol{m}$ and $\boldsymbol{n}$ are the formulas which stand for the positive integers m and n respectively, $E(n) = 1$ if $\{\boldsymbol{A}_n\}(\boldsymbol{n})$ is not convertible into one of the formulas $1, 2, 3, \cdots$, and $E(n) = m + 1$ if $\{\boldsymbol{A}_n\}(\boldsymbol{n})$ conv $\boldsymbol{m}$ and $\boldsymbol{m}$ is one of the formulas $1, 2, 3, \cdots$. The function E is effectively calculable and is there-

fore λ-definable, by a formula $\mathfrak{e}$. The formula $\mathfrak{e}$ has a normal form, since $\mathfrak{e}(1)$ has a normal form. But $\mathfrak{e}$ is not any one of the formulas $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \cdots$, because, for every n, $\mathfrak{e}(\boldsymbol{n})$ is a formula which is not convertible into $\{\boldsymbol{A}_n\}(\boldsymbol{n})$. And this contradicts the property of the enumeration $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3, \cdots$ that it contains all well-formed formulas which have a normal form.

COROLLARY 1. *The set of well-formed formulas which have no normal form is not recursively enumerable.*[24]

For, to outline the argument, the set of well-formed formulas which have a normal form is recursively enumerable, by Theorem XV. If the set of those which do not have a normal form were aslo recursively enumerable, it would be possible to tell effectively of any well-formed formula whether it had a normal form, by the process of searching through the two enumerations until it was found in one or the other. This, however, is contrary to Theorem XVIII.

This corollary gives us an example of an effectively enumerable set (the set of well-formed formulas) which is divided into two non-overlapping subsets of which one is effectively enumerable and the other not. Indeed, in view of the difficulty of attaching any reasonable meaning to the assertion that a set is enumerable but not effectively enumerable, it may even be permissible to go a step further and say that here is an example of an enumerable set which is divided into two non-overlapping subsets of which one is enumerable and the other non-enumerable.[25]

COROLLARY 2. *Let a function F of one positive integer be defined by the rule that $F(n)$ shall equal* 2 *or* 1 *according as n is or is not the Gödel representation of a formula which has a normal form. Then F (if its definition be admitted as valid at all) is an example of a non-recursive function of positive integers.*[26]

This follows at once from Theorem XVIII.

---

[24] This corollary was proposed by J. B. Rosser.

The outline of proof here given for it is open to the objection, recently called to the author's attention by Paul Bernays, that it ostensibly requires a non-constructive use of the principle of excluded middle. This objection is met by a revision of the proof, the revised proof to consist in taking any recursive enumeration of formulas which have no normal form and showing that this enumeration is not a complete enumeration of such formulas, by constructing a formula $\mathfrak{e}(\boldsymbol{n})$ such that (1) the supposition that $\mathfrak{e}(\boldsymbol{n})$ occurs in the enumeration leads to contradiction (2) the supposition that $\mathfrak{e}(\boldsymbol{n})$ has a normal form leads to contradiction.

[25] Cf. the remarks of the author in *The American Mathematical Monthly*, vol. 41 (1934), pp. 356-361.

[26] Other examples of non-recursive functions have since been obtained by S. C. Kleene in a different connection. See his forthcoming paper, "General recursive functions of natural numbers."

Consider the infinite sequence of positive integers, $F(1), F(2), F(3), \cdots$. It is impossible to specify effectively a method by which, given any n, the n-th term of this sequence could be calculated. But it is also impossible ever to select a particular term of this sequence and prove about that term that its value cannot be calculated (because of the obvious theorem that if this sequence has terms whose values cannot be calculated then the value of each of those terms 1). Therefore it is natural to raise the question whether, in spite of the fact that there is no systematic method of effectively calculating the terms of this sequence, it might not be true of each term individually that there existed a method of calculating its value. To this question perhaps the best answer is that the question itself has no meaning, on the ground that the universal quantifier which it contains is intended to express a mere infinite succession of accidents rather than anything systematic.

There is in consequence some room for doubt whether the assertion that the function F exists can be given a reasonable meaning.

THEOREM XIX. *There is no recursive function of two formulas* $\boldsymbol{A}$ *and* $\boldsymbol{B}$, *whose value is* 2 *or* 1 *according as* $\boldsymbol{A}$ *conv* $\boldsymbol{B}$ *or not.*

This follows at once from Theorem XVIII and the Lemma preceding it.

As a corollary of Theorem XIX, it follows that the Entscheidungsproblem is unsolvable in the case of any system of symbolic logic which is ω-consistent (ω-widerspruchsfrei) in the sense of Gödel (*loc. cit.*, p. 187) and is strong enough to allow certain comparatively simple methods of definition and proof. For in any such system the proposition will be expressible about two positive integers a and b that they are Gödel representations of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A}$ is immediately convertible into $\boldsymbol{B}$. Hence, utilizing the fact that a conversion is a finite sequence of immediate conversions, the proposition $\Psi(a, b)$ will be expressible that a and b are Gödel representations of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ such that $\boldsymbol{A}$ conv $\boldsymbol{B}$. Moreover if $\boldsymbol{A}$ conv $\boldsymbol{B}$, and a and b are the Gödel representations of $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively, the proposition $\Psi(a, b)$ will be provable in the system, by a proof which amounts to exhibiting, in terms of Gödel representations, a particular finite sequence of immediate conversions, leading from $\boldsymbol{A}$ to $\boldsymbol{B}$; and if $\boldsymbol{A}$ is not convertible into $\boldsymbol{B}$, the ω-consistency of the system means that $\Psi(a, b)$ will not be provable. If the Entscheidungsproblem for the system were solved, there would be a means of determining effectively of every proposition $\Psi(a, b)$ whether it was provable, and hence a means of determining effectively of every pair of formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ whether $\boldsymbol{A}$ conv $\boldsymbol{B}$, contrary to Theorem XIX.

In particular, if the system of *Principia Mathematica* be ω-consistent, its Entscheidungsproblem is unsolvable.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

ASYMPTOTIC SOLUTIONS OF CERTAIN ORDINARY DIFFERENTIAL EQUATIONS ASSOCIATED WITH MULTIPLE ROOTS OF THE CHARACTERISTIC EQUATION.

By Hugh L. Turrittin.

Introduction. A linear differential equation

$$(1)\quad \mathcal{L}(y) \equiv \frac{d^n y}{dx^n} + \rho^r P_1(x,\rho)\frac{d^{n-1}y}{dx^{n-1}} + \rho^{2r} P_2(x,\rho)\frac{d^{n-2}y}{dx^{n-2}} + \cdots + \rho^{nr} P_n(x,\rho) y =$$

with continuous coefficients that may be represented by convergent series

$$(2)\quad P_i(x,\rho) = \sum_{j=0}^{\infty} P_{i,j}(x)\rho^{-j}, \quad (i = 1, 2, \cdots, n;\ a \leqq x \leqq b;\ |\rho| > R_1).$$

has under certain well known conditions[1] a fundamental set of solutions of the form

$$(3)\qquad y_i(x,\rho) = \Big(\sum_{j=0}^{\zeta} y_{ij}(x)\rho^{-j} + \frac{\mathcal{E}(x,\rho)}{\rho^{\zeta+1}} \Big) \exp\Big[\int_a^x \Omega_i(t,\rho)\,dt\Big]$$
$$(a \leqq x \leqq b;\ |\rho| > R_3 > R_1;\ \gamma \leqq \arg.\rho \leqq \theta)$$

where $\mathcal{E}$ is a generic symbol for a uniformly bounded function. The particular form of the solutions as exhibited in (3) has been derived in the classical theory subject to the hypothesis that the characteristic equation possesses distinct roots throughout the fundamental interval (a, b). The object of the present paper is to generalize the theory and dispense in part with the above mentioned hypothesis. Schlesinger stated[2] and the analysis below will prove that, if multiple characteristic roots occur, it is necessary to modify the right-hand member of (3) so as to include fractional, as well as full, powers of the parameter ρ.

[1] The development of the theory concerned with solutions of type (3) may be traced through the works of H. Poincaré, *Acta Mathematica*, vol. 8 (1886), pp. 295-344; J. Horn, *Mathematische Annalen*, vol. 52 (1899), pp. 271-292; L. Schlesinger, *Mathematische Annalen*, vol. 63 (1907), pp. 277-300; G. D. Birkhoff, *Transactions of the American Mathematical Society*, vol. 9 (1908), pp. 219-231; O. Perron, *Sitzungsberichte der Heidelberger Akamedie der Wissenschaften, Mathematik Naturwissenschaften*, Klasse 1919 (A. 6); J. Tamarkin, *Thesis*, Petrograd (1917), and *Mathematische Zeitschrift*, vol. 27 (1928), pp. 1-54.

[2] L. Schlesinger, *op. cit.*, p. 282.

The first part of the demonstration is devoted to the construction of n formal series solutions. The existence proof proper, which then follows, begins with the study of a differential equation satisfied by n functions $z_i(x, \rho)$ created from the formal series by cutting them short at some particular term. The asymptotic relation of these functions to actual solutions of the differential equation (1) is established by means of a convenient integral equation and certain related inequalities.

The characteristic equation. Previous investigators have shown that to each simple root of the characteristic equation

$$(4) \qquad F(\phi) \equiv \phi^n + P_{1,0}(x)\phi^{n-1} + P_{2,0}(x)\phi^{n-2} + \cdot \cdot \cdot + P_{n,0}(x) = 0$$

there may be made to correspond a formal solution of differential equation (1). In order to procure n independent expansions when multiple roots are admitted, it would therefore seem necessary that with each k-fold characteristic root $\phi(x)$ there must be associated k formal series. This will essentially be done, not, however, with respect to the single characteristic equation (4), but with reference to a sequence of auxiliary characteristic equations.

Throughout the text it will be convenient to indicate derivatives by superscripts in parentheses. In particular reference will be made to the function $F^{(j)}(\phi)$, the j-th derivative with respect to ϕ of the left member of the characteristic equation (4).

Hypotheses. It will be understood hereafter that a root which is distinct from all others throughout the fundamental interval (a, b) is to be called *simple*; while the term *multiple* will be reserved for all those roots, and only those, which coincide over the entire interval (a, b). It is to be assumed that all roots, whether of equation (4), or of the subsequent auxiliary characteristic equations, are to be, in the sense just defined, either simple or multiple. This paper will not be concerned with roots of a more complicated nature.[3]

Furthermore it is presumed that the functions $P_{i,j}(x)$ in series (2) possess derivatives of all orders.

Later on it will be assumed that certain Wronskians do not vanish. Such assumptions will be indicated by the letter $(\mathfrak{A})$.

If a change of variable is made by setting $y = ve^{-\rho^r \epsilon x}$, where ϵ is a small positive constant, the general form of the differential equation (1) remains unaltered. However, the characteristic equation is affected and the char-

[3] See, however, in this regard R. E. Langer, *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 90-106.

acteristic roots are increased by an amount equal to ϵ. Such a change in the equation can always be employed to avoid zero characteristic roots. It therefore can and will be assumed without imposing an additional restriction that $P_{n,0}(x) \neq 0$.

The first change of variable and the Puiseux polygon. As an initial step toward procuring formal series solutions corresponding to a k-fold characteristic root $\phi(x)$, substitute in (1)

$$y = v \exp\left[\int_a^x \rho^r \phi(t)\, dt \right].$$

The differential equation (1) then acquires the new form

$$v^{(n)} + p_1(x,\rho) v^{(n-1)} + p_2(x,\rho) v^{(n-2)} + \cdots + p_n(x,\rho) v = 0 \tag{5}$$

where the leading term in the series expansion of the coefficient $p_i(x,\rho)$ is formally $[\rho^{ir}/(n-i)!]F^{(n-i)}(\phi)$. Due to the k-fold multiplicity of $\phi(x)$, $F^{(n-i)}(\phi) \equiv 0$ if $i = n, n-1, \cdots$ or $n-k+1$. Consequently the coefficients have the structure

$$p_i(x,\rho) = \rho^{j_i} p_{i,0}(x) + \rho^{j_i} \sum_{h=1}^{\infty} p_{i,h}(x) \rho^{-h} \qquad (i = 1, \cdots, n) \tag{6}$$

with

$$\begin{array}{lll} j_i \leqq ir & \text{for} & i = 1, 2, \cdots, n-k-1; \\ j_i = (n-k)r & \text{for} & i = n-k, \text{ and} \\ j_i \leqq ir - 1 & \text{for} & i = n-k+1, \cdots, n. \end{array}$$

Hereafter the leading term in an expansion of a function in descending powers of ρ will be called the *principal term.*

The principal terms of the series (6) can be advantageously represented by a diagram in the following manner:

Let $p_{i,0}(x)$ correspond to the point with coördinates (i, j_i); join the origin to the point $(n, 0)$ not only by a straight line-segment, but also by a broken line to form a convex polygon with vertices falling on points (i, j_i) and enclosing the greatest possible positive area. The result is an adaptation of the familiar Puiseux polygon.[4] By virtue of the construction the slope s_i of the i-th line-segment (i. e. the slope of the i-th side of the polygon counting from left to right) is less than the slope s_{i-1} of the preceding segment, and the points (i, j_i) will all fall on, or below, the broken line. The significance of the Puiseux polygon lies in the fact that there can be associated with each

[4] V. Puiseux, *Journal de Mathématiques Pures et Appliquées*, vol. 11 (1889), p. 319.

of the line-segments formal solutions of equation (1) in number equal to the length of the projection of the respective segment, or side, on the x-axis.

Case I. Formal solutions pertaining to segments with negative or zero slope are derived as follows:

Let the coördinates of the right terminus of the chain of segments with positive slopes be designated by (m, j_m). Then replace the dependent variable v in equation (5) by the series

$$v_0(x) + v_1(x)/\rho + v_2(x)/\rho^2 + \cdots.$$

Equate the coefficients of the various powers of ρ to zero; and thus derive a sequence of linear differential equations, each of order $n - m$ from which the $v_i(x)$'s may be successively determined. The first differential equation of this sequence is homogeneous, and therefore yields $n - m$ independent evaluations of $v_0(x)$. The other $v_i(x)$'s may be determined by quadratures once the $v_0(x)$'s have been found. Hence there exist for the differential equation (1) $n - m$ formal solutions of the structure

$$(7) \quad (v_{j,0}(x) + v_{j,1}(x)/\rho + \cdots) \exp[\int_a^x \rho^r \phi(x)\,dx] \quad (j = 1, 2, \cdots, n-m).$$

It will be assumed, $(\mathfrak{a})$, that the Wronskian

$$(8) \qquad W(v_{1,0}(x), \cdots, v_{n-m,0}(x)) \neq 0;$$

and the hypothesis (8) would be satisfied, it is to be noted, if $p_{m,0}(x)$ were non-vanishing on (a, b).

Case II. There remains for consideration only those segments of positive slope which lie between the points $(n - k, nr - kr)$ and (m, j_m), for the first segment pertains to characteristic roots other than $\phi(x)$. Let attention be fixed upon any particular one of these segments, the σ-th for instance, and let μ_σ denote the length of the projection of the segment upon the x-axis. To obtain formal solutions to the number μ_σ, one, or more, transformations will be necessary. Let

$$(9) \qquad v = u \exp[\int_a^x \lambda^a \psi_1(x)\,dx]$$

with $\lambda^a = \rho^{\alpha/\beta}$, where α/β is a fraction in lowest terms equal to the slope s_σ of the segment in question.

Transformation (9) modifies simultaneously the Puiseux polygon and the differential equation (5). Those segments up to and including the

$(\sigma-1)$-th remain unaltered. The chain of segments to the right of the $(\sigma-1)$-th is, however, replaced by a single segment of slope s_σ. The differential equation (5) is reduced by (9) to a form which, in so far as the last $(\mu_\sigma+1)$ terms are concerned, can be indicated as follows:

$$\cdots+q_0(x,\lambda)u^{(\mu_\sigma)}+q_1(x,\lambda)u^{(\mu_\sigma-1)}+\cdots+q_{\mu_\sigma}(x,\lambda)u=0$$

in which

$$(10)\quad q_j(x,\lambda)=\frac{\lambda^{aj}G^{(\mu_\sigma-j)}(\psi_1)}{(\mu_\sigma-j)\,!}+\lambda^{aj}\sum_{h=1}^{\infty}q_{j,h}(x)\lambda^{-h}\qquad (j=0,1,\cdots,\mu_\sigma)$$

with $G^{(\mu_\sigma-j)}(\psi)$ representing the $(\mu_\sigma-j)$-th derivative of the function

$$(11)\qquad G(\psi)\equiv\psi^{n-\mu_1-\mu_2-\cdots-\mu_\sigma}\{C_0(x)\,\psi^{\mu_\sigma}+\cdots+C_{\mu_\sigma}(x)\}=0$$

which, as is indicated, has been set equal to zero in (11) to form the auxiliary characteristic equation. The coefficients $C_0(x)$ and $C_{\mu_\sigma}(x)$ are equal to those functions $p_{i,0}(x)$ of series (6) which correspond to the vertices at the left and right ends respectively of the σ-th segment in the polygon. The intermediate coefficients of (11) likewise are equal to $p_{i,0}(x)$'s which correspond to other points (i,j_i) that fall on the σ-th segment. These facts make it clear that transformation (9) has in effect shifted the σ-th segment of the original polygon over to the extreme right of the figure.

The portion of the auxiliary characteristic equation (11) which is contained within the braces is a polygon in ψ^β. Hence there are at least β different non-zero roots of the equation (11). If one root has been found, $\beta-1$ others can be computed by multiplying the known one by the various β-th roots of unity. When in the subsequent discussion the function $\psi_1(x)$ which appears in substitution (9) is identified with a particular root of (11), it should be born in mind that β such roots can be considered simultaneously by allowing the β-th roots of unity to be absorbed into the parameter λ.

The function $\psi_1(x)$ of substitution (9), which as yet has not been specified, will be selected as a non-zero κ-fold multiple root of the auxiliary characteristic equation. Such a choice has an immediate effect upon the differential equation causing the leading terms in those series of (10) to vanish for which $j=\mu_\sigma,\mu_\sigma-1,\cdots$ and $\mu_\sigma-\kappa+1$. The effect upon the polygon is even more important to note. One of two things must take place. Either the right-hand portion of the new diagram which corresponds to the old σ-th segment is broken up into at least two smaller line-segments, or the right-hand segment is subject to a decrease in slope. The break-up certainly takes place if $\beta>1$ and the decrease in slope occurs if $\beta=1$.

Since κ formal series are to be associated with ψ_1, we need no longer be concerned with the original σ-th segment, but must concentrate upon those segments between $x = n - \kappa$ and $x = n$, all of which have slopes less than s_σ. On the one hand certain of these segments may have zero or negative slope and to such segments formal series will at once be associated by the method described in Case I, subject of course to a hypothesis ($\mathfrak{a}$) that a certain Wronskian, the counterpart of (8), is non-vanishing. On the other hand, if formal solutions are to be associated with those segments of positive slope between $x = n - \kappa$ and $x = n$, further exponential transformations must be made.

The Puiseux polygon which has undergone two changes, one brought about by substitution (9), the other by virtue of a special choice of ψ_1, supplies the data for selecting the next exponential transformation; in fact the power of the parameter which must appear under the integral sign equals the slope of the line segment to which formal solutions are eventually to be associated. For instance, if a segment between $x = n - \kappa$ and $x = n$ has a slope Δ, $(0 < \Delta < s_\sigma)$, the proper substitution is

$$u = w \exp\left[\int_a^x \rho^\Delta \chi(x)\, dx\right]. \tag{12}$$

The effect of this transformation can be readily visualized in as much as the effect of (9), a typical transformation, has already been described in detail. Transformation (12) introduces, as a rule, a new and smaller fractional power of ρ into the differential equation and the polygon is modified to the extent that the segment with slope Δ is shifted to the right side of the diagram. Then a second auxiliary characteristic equation is formed by equating to zero the principal term in the series expansion of the coefficient of w in the new linear differential equation. The function $\chi(x)$ is definitely chosen as one of the non-zero secondary auxiliary characteristic roots with the result that the Puiseux polygon is either broken into smaller segments or the slope of the right-hand segment is decreased. Formal solutions are immediately associated with all those segments which have, as a result of the reduction, zero or negative slopes. (This reapplication of Case I calls, in general, for another assumption of type $\mathfrak{a}$.) To associate formal solutions with the segments of positive slope still further transformation are required, and hence, *mutatis mutandis*, the whole process is repeated.

A succession of reductions, each involving a decrease in slope, unaccompanied by a break-up of the segments, must give rise eventually to a segment of zero slope and the process is thereby terminated (Case I, $\mathfrak{a}$). That the

zero slope is ultimately attained is made evident by reexamining the typical reduction arising from (9) when the appropriate values $\beta = 1$, $\lambda = \rho$, and the integer $\alpha \leq r - 1$ are substituted in (9). The reduction will result in a slope $\leq r - 2$ for the right-hand segment, because no new fractional power of the parameter has been introduced. A second reapplication of the process with a new exponent $\alpha_1 \leq r - 2$ will result in a slope $\leq r - 3$, etc., until with not more than $r - 1$ successive reductions, with β continually equal to one, a zero slope is attained for the right-hand segment of the polygon. On the other hand a series of reductions, accompanied by a succession of break-ups, must finally result in a short segment with unit projection on the x-axis. The auxiliary characteristic equation, which comes from a reapplication of the process to such a segment, is of the first degree and can not possess a multiple root. Therefore from this stage on only decreases in slope can occur when the proper exponential transformations are made. Finally a zero slope is attained and formal solutions procured. This reasoning establishes the fact that there exist n formal solutions for differential equation (1) of the structure

$$(13)\quad \left(z_{i,0}(x) + \frac{z_{i,1}(x)}{\rho^{1/\beta_i}} + \frac{z_{i,2}(x)}{\rho^{2/\beta_i}} + \cdots\right) \exp\left[\int_a^x \Omega_i(t,\rho)\,dt\right] \quad (i = 1, \cdots, n)$$

where $\Omega_i(t, \rho)$ is an abbreviation for the sum

$$\rho^r \Omega_{i,1}(t) + \rho^{r-1/\beta_i}\Omega_{i,2}(t) + \cdots + \rho^{1/\beta_i}\Omega_{i,r\beta_i-1}(t).$$

It is, moreover, allowable to attribute to ρ^{1/β_i} any one of its β_i values. Furthermore, if for some particular value of i, $\Omega_i(t, \rho)$ is different from all other $\Omega_j(t, \rho)$, $i \neq j$, then it follows as a specialization of (8) that

$$(14)\qquad z_{i,0}(x) \neq 0.$$

The related differential equation. In order to obtain an asymptotic relationship between the formal series (13) and actual solutions of the given differential equation (1), a study will be made of the differential equation satisfied by the functions $z_i(x, \rho)$ defined by arbitrarily terminating series (13); i. e.,

$$z_i(x, \rho) = \eta_i(x, \rho) \exp\left[\int_a^x \Omega_i(t, \rho)\,dt\right] \qquad (i = 1, \cdots, n).$$

with

$$\eta_i(x, \rho) = z_{i,0}(x) + z_{i,1}(x)/\rho^{1/\beta_i} + \cdots + z_{i,\zeta\beta_i}(x)/\rho^{\zeta}.$$

These functions $z_i(x, \rho)$ are solutions of the differential equation

$$(15)\qquad \begin{vmatrix} z^{(n)} & \cdot & \cdot & \cdot & z^{(1)} & z \\ z_1^{(n)} & \cdot & \cdot & \cdot & z_1^{(1)} & z_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ z_n^{(n)} & \cdot & \cdot & \cdot & z_n^{(1)} & z_n \end{vmatrix} = 0$$

which has been written as a determinant. This determinant becomes, when expanded in terms of the elements of the first row,

$$(16)\qquad \mathfrak{M}(z) \equiv z^{(n)} + \rho^r Q_1(x,\rho) z^{(n-1)} + \cdots + \rho^{rn} Q_n(x,\rho) z = 0$$

where the coefficients are quotients of two n-th order determinants; i. e.,

$$(17)\qquad Q_i(x,\rho) = (-1)^i A_i(x,\rho)/\rho^{ir} A_0(x,\rho) \qquad (i = 1, \cdots, n)$$

In particular $A_0(x,\rho)$ is the Wronskian $W(z, \cdots, z_n)$.

Since the various functions $z_i(x,\rho)$, which are associated with the introduction of a particular fractional power of ρ, may be permuted among themselves in (15) without essentially changing the differential equation (16), it is evident that the expansions of the determinants $A_i(x,\rho)$ proceed according to decreasing *full* powers. Furthermore these expansions are finite, not infinite series.

Structure of the coefficients $Q_i(x,\rho)$. To make certain that $Q_i(x,\rho)$ can be expanded in convergent power series in ρ^{-1}, it will be sufficient to show that $A_0(x,\rho) \exp[-\int_a^x (\sum_{i=1}^n \Omega_i)\, dt]$, which is later indicated by $\Lambda A_0(x,\rho)$, is bounded away from zero as $\rho \to \infty$. The element in the i-th row and j-th column of $A_0(x,\rho)$ is

$$d^{n-j} z_i / dx^{n-j} = (d_i/dx + \Omega_i(x,\rho))^{n-j} \eta_i(x,\rho) \; \exp[\int_a^x \Omega_i dt]$$
$$= \mathfrak{D}_i^{n-j} \eta_i(x,\rho) \exp[\int_a^x \Omega_i dt].$$

The symbolic operators $\mathfrak{D}_i = (d_i/dx + \Omega_i(x,\rho))$ and d_i/dx will be especially useful in ascertaining the character of $A_0(x,\rho)$ as $\rho \to \infty$. It should be understood, however, that, when the differential operator d_i/dx acts upon a function with subscript other than i, it is to be considered equivalent to differentiating a constant. For example this notation permits one first to think of the Wronskian as a Vandermonde determinant; i. e.,

$$\begin{vmatrix} \eta_1 & \cdots & \eta_n \\ \eta_1^{(1)} & \cdots & \eta_n^{(1)} \\ \cdot & \cdots & \cdot \\ \eta_1^{(n-1)} & \cdots & \eta_n^{(n-1)} \end{vmatrix} = \{\prod_{i>j}(d_i/dx - d_j/dx)\}\eta_1\eta_2\cdots\eta_n; \quad \begin{matrix}(i=2,3,\cdots,n;\\ j=1,\cdots,n-1),\end{matrix}$$

and secondly to write

$$\text{(18)} \quad \Lambda A_0(x,\rho) = \begin{vmatrix} \mathcal{D}_1^{n-1} & \cdots & \mathcal{D}_1 & 1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ \mathcal{D}_n^{n-1} & \cdots & \mathcal{D}_n & 1 \end{vmatrix} \eta_1\eta_2\cdots\eta_n = \{\prod_{i<j}(\mathcal{D}_i - \mathcal{D}_j)\}\eta_1\eta_2\cdots\eta_n.$$

$$(i=1,\cdots,n-1;\ j=2,\cdots,n),$$

a convenient expression for $\Lambda A_0(x,\rho)$. A glance at the right-hand member of (18) shows, if we consider the principal terms in the expansions of the various factors, that as $\rho\to\infty$ the behavior of $\Lambda A_0(x,\rho)$ is essentially the same as that of the product

$$\text{(19)} \qquad \{\prod_{i<j}(\Omega_i - \Omega_j)\}\{\prod_{i<j}(d_i/dx - d_j/dx)\}z_{1,0}z_{2,0}\cdots z_{n,0}$$

where all those combinations of i and j $(i<j)$, for which $\Omega_i - \Omega_j \neq 0$, are to be used in the first brace, and in the second all other combinations of $i<j\leq n$ are to be used. The product (19) and likewise $\Lambda A_0(x,\rho)$ are bounded away from zero for $|\rho|$ sufficiently large due to the fact that in (19) only non-vanishing factors occur, some of these factors are of type $(\Omega_i - \Omega_j)$, some of type (14), and others are non-zero Wronskians of type (8). Not only does the above argument establish the fact that $Q_i(x,\rho)$ can be expanded in convergent series in descending powers of ρ, but it also shows that the functions $z_1(x,\rho),\cdots,z_n(x,\rho)$ are independent.

To demonstrate that $Q_i(x,\rho)$ is uniformly bounded in the neighborhood of $\rho=\infty$, it is enough to show that

$$\text{(20)} \qquad Q_i(x,\rho) = \sum_{h=0}^{\infty} Q_{i,h}(x)\rho^{-h} \qquad (i=1,\cdots,n).$$

Let $E_j(\mathcal{D}_1,\cdots,\mathcal{D}_n)$ denote the j-th elementary symmetric function $\Sigma\mathcal{D}_1\mathcal{D}_2\cdots\mathcal{D}_j$ and $E_0=1$; and then, from the fact [5] that the determinant

[5] E. R. Heineman, *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 464-476.

$$\begin{vmatrix} \mathcal{D}_1^{n} & \mathcal{D}_1^{n-1} & \cdots & \mathcal{D}_1^{j+1} & \mathcal{D}_1^{j-1} & \cdots & \mathcal{D}_1 & 1 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot & \cdot \\ \mathcal{D}_n^{n} & \mathcal{D}_n^{n-1} & \cdots & \mathcal{D}_n^{j+1} & \mathcal{D}_n^{j-1} & \cdots & \mathcal{D}_n & 1 \end{vmatrix} = E_{n-j}(\mathcal{D}_1,\cdots,\mathcal{D}_n)\prod_{i<h}(\mathcal{D}_i-\mathcal{D}_h),$$

$$(j=0,1,\cdots,n;\ i=1,\cdots,n-1;\ h=2,\cdots,n),$$

it follows that

$$(21)\qquad \Lambda A_i(x,\rho)=E_i(\mathcal{D}_1,\cdots,\mathcal{D}_n)\prod_{j<h}(\mathcal{D}_j-\mathcal{D}_h)\eta_1\eta_2\cdots\eta_n.$$
$$(i=1,\cdots,n;\ j=1,\cdots,n-1;\ h=2,\cdots,n).$$

If it is taken into account that the power of ρ in the principal term of $E_i(\mathcal{D}_1,\cdots,\mathcal{D}_n)$ can not exceed ρ^{ir} and then if formulae (18) and (21) are compared, it will be evident that the quotients (17) must have the structure (20).

A comparison of the related and original differential equation. The discussion of the last section brings out the structural similarity of the given differential equation (1) and the related equation (16). The resemblance will be made more striking by showing (see inequality (24)) that the respective coefficients of these differential equations are to a certain extent identical; i. e.,

$$P_{i,j}(x)\equiv Q_{i,j}(x)\quad\text{for}\quad \begin{cases} i=1,\cdots,n \\ j=0,1,\cdots,\zeta+r. \end{cases}$$

With this end in view consider the n differences

$$(22)\qquad \mathfrak{M}(z_i)-\mathcal{L}(z_i).$$

Since $z_i(x,\rho)$ is a solution of (16)

$$\mathfrak{M}(z_i(x,\rho))\equiv 0.$$

The substitution of $z_i(x,\rho)$ in (1) must cause a number of terms to vanish by virtue of the fact that $z_i(x,\rho)$ is a portion of a formal series solution; those terms which do not in general vanish may be indicated as follows:

$$(23)\quad \mathcal{L}(z_i(x,\rho))\equiv\rho^{\tau_i-\zeta-1/\beta_i}\,(f_{i,0}(x)+f_{i,1}(x)/\rho^{1/\beta_i}+\cdots)\exp[\int_a^x\Omega_i dt]$$
$$(i=1,\cdots,n).$$

The f's are appropriate functions of x, and τ_i designates the power of ρ occurring in the principal term of the expansion of $\prod_j(\Omega_j-\Omega_i)$, where $j\leq n$

ranges over all values such that $\Omega_j \neq \Omega_i$. The differences (22) may be rewritten by utilizing (23) as

$$\sum_{j=1}^{n} \rho^{jr}(P_j - Q_j) z_i^{(n-j)} \equiv \mathcal{L}(z_i) \qquad (i = 1, \cdots, n).$$

These identities are treated as n simultaneous equations defining the unknowns $(P_1 - Q_1), \cdots, (P_n - Q_n)$. Each of these unknowns can be expressed by means of Cramer's rule as a quotient of two determinants. The symbolic operators $\mathcal{D}_i$ are used to determine the nature of these quotients, just as was done for quotients (17), and it is found that

$$(P_i(x,\rho) - Q_i(x,\rho)) = \rho^{-\zeta-r-1} \sum_{j=0}^{\infty} g_{ij}(x)\rho^{-j} \qquad (i = 1, \cdots, n).$$

A sufficient number of hypotheses have been made to assure continuity in the functions under consideration; hence it is possible to select the positive constants S and R_2 sufficiently large so that

$$(24) \qquad | P_i(x,\rho) - Q_i(x,\rho) | < S/\rho^{\zeta+r+1}$$
$$(i = 1, \cdots, n;\ |\rho| > R_2 > R_1;\ a \leq x \leq b).$$

Asymptotic approximations to solutions. To determine the way in which the functions $z_i(x,\rho)$ approximate solutions of (1), the original differential equation (1) is first rewritten

$$\mathcal{M}(y) = \mathcal{N}(y) \quad \text{with} \quad \mathcal{N}(y) \equiv \mathcal{M}(y) - \mathcal{L}(y)$$

and this formally non-homogeneous differential equation is then replaced by the equivalent integral equation [6]

$$(25) \quad y(x,\rho) = \sum_{i=1}^{n} c_i z_i(x,\rho) + \int_a^x \left(\sum_{i=1}^{n} z_i(x,\rho) Z_i(\xi,\rho) \mathcal{N}(y(\xi,\rho)) \right) d\xi.$$

Here the c_i's are arbitrary constants and the adjoint functions $Z_i(x,\rho)$ are defined by the system of equations

$$\sum_{i=1}^{n} z_i^{(j)}(x,\rho) Z_i(x,\rho) = \begin{cases} 0 & \text{for} \quad j = 0, 1, \cdots, n-2 \\ 1 & \phantom{\text{for}} \quad j = n-1. \end{cases}$$

It is readily deduced, after applying Cramer's rule, that the adjoint functions have the structure

$$Z_i(x,\rho) = \rho^{-\tau_i} H_i(x,\rho) \exp\left[-\int_a^x \Omega_i dt\right] \qquad (i = 1, \cdots, n).$$

---

[6] Cf. L. Schlesinger, *Handbuch der Theorie der linearen Differentialgleichungen*, vol. 1, p. 78.

The constants τ_i are the same as those appearing in identities (23) and the symbols $H_i(x, \rho)$ designate uniformly bounded functions.

A positive constant H may be selected sufficiently large so that

$$|H_i(x,\rho)| < H \qquad (a \leq x \leq b;\ |\rho| > R_2),$$

and also

$$|\mathcal{D}_i{}^j \eta_i(x,\rho)| < H\,|\rho^{ir}| \qquad (i = 1, \cdots, n;\ j = 0, 1, \cdots, n-1).$$

Hence

$$(26) \qquad |z_i{}^{(j)}(x,\rho) Z_i(\xi,\rho)| < H^2\,|\rho^{jr-\tau_i} \exp[\int_\xi^x \Omega_i dt]| \qquad (i = 1, \cdots, n;\ j = 0, 1, \cdots, n-1).$$

A region defined by the inequalities

$$a \leq x \leq b, \quad \gamma \leq \arg\rho \leq \theta, \quad \text{and} \quad |\rho| > R_3 > R_2$$

will be called a Ξ-region, if it is possible to rearrange the subscripts on the Ω_i's in such a way that the real portions of the functions $\Omega_i(x,\rho)$ can be ordered as follows:

$$\mathcal{R}(\Omega_i(x,\rho)) < \mathcal{R}(\Omega_{i+1}(x,\rho))$$
$$\text{unless } \Omega_i(x,\rho) \equiv \Omega_{i+1}(x,\rho), \text{ and then } \mathcal{R}(\Omega_i) = \mathcal{R}(\Omega_{i+1}) \qquad (i = 1, 2, \cdots, n-1).$$

Within such a region

$$(27) \quad |\exp[\int_\xi^x \Omega_1 dt]| \leq |\exp[\int_\xi^x \Omega_2 dt]| \leq \cdots \leq |\exp[\int_\xi^x \Omega_i dt]| \leq \cdots \leq |\exp[\int_\xi^x \Omega_n dt]|$$

for $x \geq \xi$.

The integral equation (25) and inequalities (24), (26), and (27) all find their counterpart in a paper by Birkhoff.[7] If we follow, therefore, Birkhoff's line of reasoning, we conclude that the integral equations

$$y_i(x,\rho) = z_i(x,\rho) + \int_a^x \Big(\sum_{h=1}^{i} z_h(x,\rho) Z_h(\xi,\rho)\Big) \mathcal{N}(y_i(\xi,\rho))\, d\xi + \int_b^x \Big(\sum_{h=i+1}^{n} z_h(x,\rho) Z_h(\xi,\rho)\Big) \mathcal{N}(y_i(\xi,\rho))\, d\xi \qquad (i = 1, \cdots, n)$$

define a fundamental set of solutions for the original differential equation (1) with the property that within a Ξ-region

[7] G. D. Birkhoff, *op. cit.*, formulas (32), (29), (43), and (7) respectively.

$$(28)\qquad y_i^{(j)}(x,\rho)=z_i^{(j)}(x,\rho)+\frac{\mathcal{E}(x,\rho)\,\exp[\int_a^x\Omega_i(t,\rho)\,dt]}{\rho^{\zeta+r+1+\tau-r(n+j)}}$$

$(a\leq x\leq b;\ |\rho|>R_3;\ i=1,\cdots,n;\ j=0,1,\cdots,n-1;\ \zeta>2rn-2r-1)$.

The constant τ is the smallest of the set $\tau_1,\tau_2,\cdots,\tau_n$. It is relation (28) that is to replace (3) when multiple roots of the characteristic equation (4) are admitted. The functions $z_i(x,\rho)$ are therefore asymptotic approximations to the solutions $y_i(x,\rho)$.

The solutions indicated in (28) are dependent upon the integer ζ. This integer is fixed, it will be recalled, by the point at which the formal series are cut short. Tamarkin and Besikowitsch [8] have shown how a solution y_n can be chosen to satisfy a relation of type (28) independently of the value of ζ. This solution is then used to reduce the order of the given differential equation to $n-1$ and through mathematical induction it is found,[9] provided the characteristic roots are unequal, that a set of n solutions exist satisfying (28) independently of ζ. This line of reasoning can readily be adopted to the present situation in which multiple roots are admitted. It is possible, for instance, that $\Omega_{n-m+1},\Omega_{n-m+2},\cdots$, and Ω_n $(1\leq m\leq n)$ are all identical. In this event m solutions $y_{n-m+1},y_{n-m+2},\cdots,y_n$ would exist which satisfy (28) independently of ζ. The induction would proceed, therefore, even more rapidly than in the classical case, for the m solutions $y_{n-m+1},\cdots,y_n$ would be used to reduce the order of the differential equation (1) to $n-m$. Hence it is concluded that there exist n independent solutions of the original differential equation (1) which have within a Ξ-region the asymptotic expansions

$$y_i^{(j)}\sim\mathcal{D}_i^{\,j}(z_{i,0}(x)+z_{i,1}(x)/\rho^{1/\beta_i}+\cdots\text{ to inf.})\exp[\int_a^x\Omega_i(t,\rho)\,dt]$$
$$(i=1,\cdots,n;\ j=0,1,\cdots,n-1)$$

where $\mathcal{D}_i^{\,j}$ represents, as before, the symbolic operator $(d_i/dx+\Omega_i(x,\rho))$ here applied j times in succession; and where the integers β_i, and the functions Ω_i and $z_{i,j}$ are the same as those appearing in the formal solutions.

UNIVERSITY OF WISCONSIN.

[8] J. Tamarkin and A. Besikowitsch, *Mathematische Zeitschrift*, vol. 21 (1924), pp. 119-123.

[9] O. Perron, *op. cit.*, pp. 14-21.

BIHARMONIC FUNCTIONS AND CERTAIN GENERALIZATIONS.[1]

By Edward Kasner.

In the following paper we study certain new characterizations of biharmonic functions in terms of conformal transformations; and certain generalizations connected with various subgroups of the total conformal group, and also with various subsets which are not groups. For example we deal not only with the six parameter inversion group, but also with the set of ∞^3 pure inversions. The main results are given in Theorems 1-8. I wish to acknowledge the valuable assistance of George Comenetz in writing this paper, especially in the proof of Theorem 5.

The Poincaré Equations.

1. Poincaré applied the term biharmonic to those functions of four real variables $F(x, y, x_1, y_1)$ which can be regarded as the real part of an analytic function of two complex variables $z = x + iy$, $z_1 = x_1 + iy_1$. The same word biharmonic is also in use to denote those functions of two real variables $f(x, y)$ which satisfy the double Laplace equation $\Delta\Delta f = 0$, of the fourth order (connected with the problem of the bending of a flat plate);[2] but the two meanings have nothing to do with each other. Biharmonic is used in this paper exclusively in the sense of Poincaré.

If $F(x, y, x_1, y_1)$ is biharmonic, then a conjugate function $G(x, y, x_1, y_1)$ exists such that $F + iG$ is an analytic function of the complex variables $x + iy$ and $x_1 + iy_1$. Hence F and G satisfy the Cauchy-Riemann differential equations with respect to each of the two complex variables; thus we have four equations of the first order:

$$(1) \qquad F_x = G_y, \quad F_y = -G_x; \qquad F_{x_1} = G_{y_1}, \quad F_{y_1} = -G_{x_1}.$$

If we differentiate these relations partially with respect to x, y, x_1 and y_1, and eliminate G, we find that F must satisfy a system of four partial differential equations of second order:

$$(2) \qquad \begin{aligned} P_1F &\equiv F_{xx} + F_{yy} = 0, \qquad & P_2F &\equiv F_{xx_1} + F_{yy_1} = 0,\\ P_3F &\equiv F_{xy_1} - F_{yx_1} = 0, \qquad & P_4F &\equiv F_{x_1x_1} + F_{y_1y_1} = 0. \end{aligned}$$

(The symbols P_1, P_2, P_3, P_4 can be thought of as differential operators, which

[1] Abstracts in *Bulletin of the American Mathematical Society*, vol. 35 (1929), pp. 290, 444.

[2] Love, *Elasticity*, 4th ed., p. 135.

we may call the Poincaré operators. In fact, P_1 is identical with the Laplace operator Δ, taken with respect to the variables x and y, and P_4 is Δ with respect to x_1, y_1.) Since

$$(3) \qquad P_1F + P_4F \equiv \partial^2F/\partial x^2 + \partial^2F/\partial y^2 + \partial^2F/\partial x_1^2 + \partial^2F/\partial y_1^2 = 0,$$

we see that a biharmonic function is necessarily harmonic in the four variables x, y, x_1, y_1; but the converse is not true: a harmonic function of x, y, x_1, y_1 need not be the real part of any analytic function of $x + iy$ and $x_1 + iy_1$ (in contrast to the case of one complex variable), for equations (2) do not follow from (3). Equations (2) are sufficient as well as necessary for F to be biharmonic.[3] They were first given by Poincaré.[4]

The Conformal Group.

2. If in an analytic function of two complex variables z and z_1, we replace z_1 by an analytic function of z, the result is an analytic function of z alone. If we consider the real part of the function of z and z_1, we see that this remark implies the following theorem on biharmonic functions:

Let

$$(4) \qquad T: \quad x_1 = \phi(x, y), \; y_1 = \psi(x, y)$$

be a substitution replacing x_1 and y_1 by functions of x and y. Then $F(x, y, x_1, y_1)$ reduces to a function K depending only on x and y:

$$(5) \qquad F(x, y, x_1, y_1) = F[x, y, \phi(x, y), \psi(x, y)] \equiv K(x, y).$$

Now replacing z_1 by an analytic function of z corresponds to using a *conformal* substitution T. Hence the obvious theorem we obtain is this:

Theorem 1. *If a conformal substitution $x_1 = \phi(x, y)$, $y_1 = \psi(x, y)$ is performed on a biharmonic function $F(x, y, x_1, y_1)$, then F becomes a harmonic function $K(x, y)$.*

We now give two fundamental converses of this theorem:

Theorem 2. *Biharmonic functions are the only functions of four real variables which are converted into harmonic functions by* EVERY *conformal substitution.*

---

[3] Osgood, *Funktionentheorie* II, 1, p. 22. (The region of definition of F must satisfy a certain connectivity restriction.)

[4] *Comptes Rendus*, vol. 96 (1883), p. 238; *Acta Mathematica*, vol. 2 (1883), p. 99; vol. 22 (1898), p. 112.

THEOREM 3. *The only substitutions which convert* ALL *biharmonic functions into harmonic functions are the conformal substitutions.*

To prove Theorems 2 and 3 we need the expression for $K_{xx}+K_{yy}$, or ΔK, in terms of F, ϕ, and ψ. This is found by differentiating (5) partially with respect to x and y (for example, $K_x = F_x + F_{x_1}\phi_x + F_{y_1}\psi_x$, etc.). The result is found to be

$$
\begin{aligned}
(6)\quad \Delta K = {} & F_{xx} + F_{yy} + 2F_{xx_1}\phi_x + 2F_{xy_1}\psi_x + 2F_{yx_1}\phi_y + 2F_{yy_1}\psi_y + F_{x_1x_1}(\phi_x^2+\phi_y^2) \\
& + 2F_{x_1y_1}(\phi_x\psi_x + \phi_y\psi_y) + F_{y_1y_1}(\psi_x^2+\psi_y^2) + F_{x_1}\Delta\phi + F_{y_1}\Delta\psi.
\end{aligned}
$$

In the case of a conformal substitution T, the conformality conditions $\phi_x = \psi_y$, $\phi_y = -\psi_x$ can be used to simplify (6) by the elimination of the derivatives. of ψ. When T is conformal, (6) in fact reduces to *our principal formula*

$$(7)\qquad \Delta K = P_1F + 2\phi_xP_2F - 2\phi_yP_3F + (\phi_x^2+\phi_y^2)P_4F.$$

Now in Theorem 2 we suppose that F is a function such that K becomes harmonic no matter what conformal substitution T is used. That is,

$$(8)\qquad [P_1 + 2\phi_xP_2 - 2\phi_yP_3 + (\phi_x^2+\phi_y^2)P_4]F(x,y,\phi,\psi) = 0$$

identically in x and y, for every conformal ϕ, ψ. We must show that F is biharmonic.

Since (8) holds for any conformal T, we may use, in particular, linear conformal substitutions, or similitudes:

$$(9)\qquad x_1 = ax - by + c, \qquad y_1 = bx + ay + d,$$

(where a, b, c, d are arbitrary constants). Then (8) becomes an identity in the six independent variables x, y, a, b, c, d. In place of c and d, we may introduce x_1 and y_1 as independent variables by solving equations (9) for c, d in terms of x, y, a, b, x_1, y_1. Then (8) becomes

$$(10)\qquad [P_1 + 2aP_2 + 2bP_3 + (a^2+b^2)P_4]F(x,y,x_1,y_1) = 0,$$

an identity in x, y, x_1, y_1, a, b. Varying a and b, we see that $P_1F, \cdots, P_4F$ must vanish at each fixed point x, y, x_1, y_1 in the region of definition of F. Hence F must be biharmonic. This proves Theorem 2.

For Theorem 3, we suppose that T is a substitution which turns every biharmonic F into a harmonic K. We must show that T is conformal.

Since the right member of (6) vanishes for any biharmonic F, it vanishes in particular when F is linear in x_1 and y_1 with constant coefficients (any linear

function of x, y, x_1, y_1 is of course biharmonic). In that case, all the terms of (6) drop out except the last two; and the coefficients of these two terms, F_{x_1} and F_{y_1}, reduce to arbitrary constants. This shows that $\Delta\phi = \Delta\psi = 0$. We next employ the particular biharmonic functions $F = a(xx_1 - yy_1) - b(xy_1 + yx_1)$, where a and b are arbitrary constants (the real part of $(a + bi)zz_1$). Substituting in (6), we have

$$2a(\phi_x - \psi_y) - 2b(\phi_y + \psi_x) = 0. \tag{11}$$

Hence $\phi_x - \psi_y = \phi_y + \psi_x = 0$, so that T must be conformal. This proves Theorem 3.

Subgroups and Subsets.

3. A function $F(x, y, x_1, y_1)$ which is made harmonic by every conformal substitution T is necessarily biharmonic, by Theorem 2. The proof of this theorem shows, in fact, that F must actually be biharmonic even if we merely require F to be converted into a harmonic function by the subgroup of the conformal group consisting of all similitudes.

Theorem 2′. *The only functions of four real variables which are converted into harmonic functions by every similitude are the biharmonic functions.*

If other *subgroups* or even *subsets which are not groups*, of the conformal group are used, however, we may expect that other functions of four variables, besides the biharmonic, will satisfy the requirement of becoming harmonic under every substitution of the subset. Larger classes of functions will thus be obtained, which we may then regard as generalizations of the biharmonic class.

We consider this problem in the case of the following subsets. (For completeness, we include similitudes in the list. The equations of substitution are given both in the complex form and in the equivalent real form (4). The letters a, b, c, d, θ stand for arbitrary real constants; and $\lambda = a + ib$, and $\mu = c + id$.)

(I) The group of ∞^4 similitudes: $z_1 = \lambda z + \mu$, or equations (9).

(II) The group of ∞^3 magnifications: $z_1 = az + \mu$, or

$$x_1 = ax + c, \qquad y_1 = ay + d. \tag{12}$$

(III) The group of ∞^3 rigid motions: $z_1 = e^{i\theta}z + \mu$, or

$$x_1 = x\cos\theta - y\sin\theta + c, \qquad y_1 = x\sin\theta + y\cos\theta + d. \tag{13}$$

(IV) The group of ∞^2 translations: $z_1 = z + \mu$, or

$$(14) \qquad x_1 = x + c, \qquad y_1 = y + d.$$

(V) The set of ∞^2 central symmetries: $z_1 = -z + \mu$, or

$$(15) \qquad x_1 = -x + c, \qquad y_1 = -y + d.$$

(VI) The group of $2\infty^2$ translations and central symmetries: $z_1 = \pm z + \mu$, or

$$(16) \qquad x_1 = \pm x + c, \qquad y_1 = \pm y + d.$$

We state the answers obtained in these six cases in a single theorem (the symbol bih stands for an arbitrary biharmonic function, the symbol $\Re$ for the real part of an analytic function of complex variables, c for an arbitrary constant; $\bar{z} = x - iy$ and $\bar{z}_1 = x_1 - iy_1$).

THEOREM 4. *The class of functions* $F(x, y, x_1, y_1)$ *which are converted into harmonic functions of* x *and* y *by all the substitutions* $x_1 = \phi(x, y)$, $y_1 = \psi(x, y)$ *of a subset of the conformal group is, in the cases of the subsets* I *to* VI, *the following:*

(I) (*Similitudes*): $F = bih(x, y, x_1, y_1)$.

(II) (*Magnifications*): $F = bih(x, y, x_1, y_1) + c(xy_1 - yx_1)$.

(III) (*Motions*): $F = bih(x, y, x_1, y_1) + c(x^2 + y^2 - x_1^2 - y_1^2)$.

(IV) (*Translations*): $F = g(x + x_1, x - x_1, y + y_1, y - y_1)$, *where* g *is harmonic in the pair of variables* $x + x_1$, $y + y_1$, *and arbitrary in* $x - x_1$, $y - y_1$; *or* $F = \Re[f(z, z_1, \bar{z} - \bar{z}_1)]$.

(V) (*Central symmetries*): $F = g(x + x_1, x - x_1, y + y_1, y - y_1)$, *where* g *is harmonic in* $x - x_1$, $y - y_1$, *and arbitrary in* $x + x_1$, $y + y_1$; *or* $F = \Re[f(z, z_1, \bar{z} + \bar{z}_1)]$.

(VI) (*Translations and central symmetries*):

$$F = g(x + x_1, x - x_1, y + y_1, y - y_1),$$

where g *is harmonic in the pairs* $x + x_1$, $y + y_1$, *and* $x - x_1$, $y - y_1$; *or*

$$F = \Re[f_1(z, z_1) + f_2(z + z_1, \bar{z} - \bar{z}_1)];$$

or

$$F = bih_1(x, y, x_1, y_1) + bih_2(x, y_1, x_1, y).[5]$$

---

[5] Two other cases may be mentioned: the group of ∞^2 similitudes leaving the origin fixed, $z_1 = \lambda z$ which gives $F = \Re[f(z, z_1, \bar{z}/\bar{z}_1)]$; and the set of ∞^2 reciprocations $z_1 = \lambda/z$ which gives $F = \Re[f(z, z_1, \bar{z}\bar{z}_1)]$.

Proof of Theorem 4.

Case I has already been settled in Theorem 2′.

To prove Theorem 4 in cases II to VI we begin as in the proof of Theorem 2 (i. e., of case I). Equation (8) must hold identically in x and y, for every ϕ, ψ of types (12) to (16) respectively. Substituting (12)-(16) into (8), we obtain identities in x, y and some of the letters a, b, c, d, θ. If we solve (12)-(16) for c and d in terms of x_1, y_1, and the other letters, we can eliminate c and d from these identities, introducing x_1 and y_1 as independent variables instead. Now if we think of x, y, x_1, y_1 as fixed, and vary a, b, θ, we find that the conditions which F must satisfy are, in the various cases, as follows:

(II) $P_1F = P_2F = P_4F = 0.$
(III) $(P_1 + P_4)F = P_2F = P_3F = 0.$
(IV) $(P_1 + 2P_2 + P_4)F = 0.$
(V) $(P_1 - 2P_2 + P_4)F = 0.$
(VI) $(P_1 + P_4)F = P_2F = 0.$

These can be looked on as systems of partial differential equations of second order, to be solved for the unknown F. Fortunately all these equations are linear, with constant coefficients, and can be solved explicitly.

The solution, in cases II and III, can be carried out most conveniently with the aid of the general identities

$$(17)\qquad \begin{aligned} (P_1)_{x_1} &= (P_2)_x - (P_3)_y; & (P_4)_x &= (P_2)_{x_1} + (P_3)_{y_1}, \\ (P_1)_{y_1} &= (P_2)_y + (P_3)_x, & (P_4)_y &= (P_2)_{y_1} - (P_3)_{x_1}. \end{aligned}$$

These are easy to verify from the definitions of $P_1, \cdots, P_4$ in (2).

First, under the conditions of case II, we see from (17) that P_3F is independent of x, y, x_1 and y_1. Hence $P_1F = P_2F = P_4F = 0$, $P_3F = 2c$ (an arbitrary constant). A particular solution of this non-homogeneous system is: $F = c(xy_1 - yx_1)$. The general solution of the corresponding homogeneous system $P_1F = P_2F = P_3F = P_4F = 0$ is: $F =$ *any biharmonic function.* Hence the general solution of the equations of case II is:

$$F = bih(x, y, x_1, y_1) + c(xy_1 - yx_1),$$

which is the result given in Theorem 4.

In case III we find from (17) that P_1F depends only on x and y, while P_4F depends only on x_1 and y_1. Since $P_1F + P_4F = 0$, in view of the different variables in P_1F and P_4F we must have $P_1F = -P_4F = 4c$. The general solution of the non-homogeneous system of equations is now obtained just as in case II.

The equations of cases IV, V, VI are simplified by the introduction of the new variables

$$(18)\qquad \begin{aligned} \xi &= x + x_1, & \eta &= y + y_1, \\ \xi_1 &= x - x_1, & \eta_1 &= y - y_1. \end{aligned}$$

Transforming the equations to these new variables, we find for IV: $F_{\xi\xi} + F_{\eta\eta} = 0$, and for V: $F_{\xi_1\xi_1} + F_{\eta_1\eta_1} = 0$. Hence in case IV, F is harmonic in ξ and η, and in case V, F is harmonic in ξ_1 and η_1, in accordance with Theorem 4.

The first statement under case VI in Theorem 4 follows from the results in cases IV and V. Proofs of the remaining statements in Theorem 4 are omitted. We remark that the work is simplified by the use of the minimal coördinates z, $\bar{z}$, z_1, $\bar{z}_1$ instead of x, y, x_1, y_1, F then being assumed analytic (see the proof of Theorem 5 below). In the proofs above it was sufficient to assume the existence of continuous third derivatives of F. (Actually all the results are valid without assuming analyticity.)

Anti-biharmonic Functions and Reverse Conformal Transformations.

4. The real part of an analytic function of z and z_1 is, by definition, biharmonic; and the real part of an analytic function of the conjugate complex variables $\bar{z}$ and $\bar{z}_1$ is also biharmonic, since in the latter case equations (1) hold with G replaced by $-G$, and then equations (2) still follow. But the real part of an analytic function of z and $\bar{z}_1$, or of $\bar{z}$ and z_1, is in general not biharmonic; we shall term it "anti-biharmonic." The sign of G is changed in only one pair of equations (1), and then instead of (2) we find

$$(19)\qquad \begin{aligned} P_1F &\equiv F_{xx} + F_{yy} = 0, & P'_2F &\equiv F_{xx_1} - F_{yy_1} = 0, \\ P'_3F &\equiv F_{xy_1} + F_{yx_1} = 0, & P_4F &\equiv F_{x_1x_1} + F_{y_1y_1} = 0 \end{aligned}$$

as the *defining equations of anti-biharmonic functions.*

A "reverse conformal" substitution T is one obtained by setting z_1 equal to an analytic function of $\bar{z}$. It is easy to modify the proofs of Theorems 1, 2, and 3 so as to show that the same theorems still hold with "anti-biharmonic" and "reverse conformal" in place of "biharmonic" and "conformal," respectively.

The question considered in Theorem 4 can also be raised here: given a set of reverse conformal substitutions, to determine the functions of x, y, x_1, y_1 which are turned into harmonic functions of x, y by every substitution of the set. Theorem 5 answers this question for the following important sets of substitutions:

(VII) The total set of ∞^2 line symmetries: $z_1 = -\bar{z}\,e^{2i\theta} + ae^{i\theta}$, or

$$(20) \qquad \begin{aligned} x_1 &= -x\cos 2\theta - y\sin 2\theta + a\cos\theta, \\ y_1 &= -x\sin 2\theta + y\cos 2\theta + a\sin\theta. \end{aligned}$$

(VIII) The total set of ∞^3 inversions: $z_1 = \dfrac{r^2}{\bar{z}-\bar{\lambda}} + \lambda$, or

$$(21) \qquad \begin{aligned} x_1 &= \frac{r^2(x-a)}{(x-a)^2+(y-b)^2} + a, \\ y_1 &= \frac{r^2(y-b)}{(x-a)^2+(y-b)^2} + b. \end{aligned}$$

(Equations (20) represent a reflection in the line whose normal makes an angle θ with the x-axis, and which is at the distance $a/2$ from the origin; (21) are the equations of an inversion in the circle of radius r, with center at the point $\lambda \equiv a + ib$. This can be seen directly from the complex forms of (20) and (21)).

SYMMETRIES AND INVERSIONS.

THEOREM 5. *The class of functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every one of a set of reverse conformal substitutions is, in the case of sets* VII *and* VIII, *the following:*

(VII) (*Line symmetries*) : $F = \Re\left[f\left(z, \bar{z}_1, \dfrac{z-z_1}{\bar{z}-\bar{z}_1}\right)\right]$.

(VIII) (*Inversions*) : $F = \Re\left[f(z, \bar{z}_1) + \text{const.}\cdot\log\dfrac{z-z_1}{\bar{z}-\bar{z}_1}\right]$,

that is, $F = \textit{antibih}(x, y, x_1, y_1) + c\tan^{-1}\dfrac{y-y_1}{x-x_1}$.[6]

These two problems seem important on account of their possible connection with automorphic functions. It will be noticed that in both cases we are dealing with subsets which do not have the group property. The analytic work is complicated since it turns out that we must solve linear partial differential equations with variable coefficients.

To prove Theorem 5 we first derive the condition, analogous to (8), for a given function $F(x, y, x_1, y_1)$ to be reduced to a harmonic function of x and y by a *reverse* conformal substitution. This is found from (6) by using the reverse conformality conditions, $\phi_x = -\psi_y$, $\phi_y = \psi_x$. The result is

$$(22) \qquad P_1F + 2\phi_x P'_2F + 2\phi_y P'_3F + (\phi_x^2 + \phi_y^2)P_4F = 0.$$

[6] The functions which are made harmonic by line symmetries include those made harmonic by inversions. This is because line symmetries are the limiting case of inversions, when the circle of inversion becomes a straight line.

Now following the procedure explained in the proof of Theorem 2, we substitute into (22) the special ϕ's of (20) and (21), and then eliminate a, b, θ by solving (20) and (21) for these parameters. In case VII we find the single equation

$$(23) \quad [(x-x_1)^2+(y-y_1)^2](P_1+P_4)F-2[(x-x_1)^2-(y-y_1)^2]P'_2F \\ -4(x-x_1)(y-y_1)P'_3F=0.$$

In case VIII, by varying the remaining parameter r we arrive at the system of three equations

$$(24) \quad P_1F=P_4F=[(x-x_1)^2-(y-y_1)^2]P'_2F \\ +2(x-x_1)(y-y_1)P'_3F=0.$$

As before, these are all linear partial differential equations of second order in the unknown $F(x,y,x_1,y_1)$, but the coefficients are no longer constants. However, it is still possible to integrate the equations explicitly. The first step is to simplify them by introducing the minimal variables

$$(25) \quad u=x+iy, \quad v=x-iy, \quad u_1=x_1+iy_1, \quad v_1=x_1-iy_1 \text{ [7]}$$

(the notation u, v, u_1, v_1 is more convenient than $z, \bar{z}, z_1, \bar{z}_1$). When this change of variables is carried out ($P_1F\equiv F_{xx}+F_{yy}=4F_{uv}$, etc.), the equations assume the forms

$$(26) \quad \text{(VII)} \quad (u-u_1)^2F_{uu_1}-(u-u_1)(v-v_1)(F_{uv}+F_{u_1v_1})+(v-v_1)^2F_{vv_1}=0,$$

$$(27) \quad \text{(VIII)} \quad F_{uv}=F_{u_1v_1}=(u-u_1)^2F_{uu_1}+(v-v_1)^2F_{vv_1}=0.$$

Now in case VII we find that equation (26) can be factored symbolically as follows:

$$(28) \quad \left[(u-u_1)\frac{\partial}{\partial u}-(v-v_1)\frac{\partial}{\partial v_1}-1\right]\left[(u-u_1)\frac{\partial}{\partial u_1}-(v-v_1)\frac{\partial}{\partial v}\right]F=0.$$

Hence if we denote by E the function obtained by applying to F the operator in the second pair of brackets, we may replace (28) by the pair of linear equations

$$(29) \quad (u-u_1)F_{u_1}-(v-v_1)F_v=E,$$

$$(30) \quad (u-u_1)E_u-(v-v_1)E_{v_1}-E=0.$$

The second of these is easily solved, with the result

---

[7] I. e., F is assumed analytic in the four real variables x, y, x_1, y_1, and this analytic function is extended to complex values of x, y, x_1, y_1. Thus F stands for a function of complex variables from this point up to the last step of the proof, when x, y, x_1, y_1 are specialized to real values again.

$$(31) \qquad E = (u - u_1) f_1 \left(v, u_1, \frac{v - v_1}{u - u_1} \right),$$

where f_1 is an arbitrary function of three variables. Substituting (31) into (29) and solving this in turn (for instance, with the aid of $\frac{v - v_1}{u - u_1}$ as an independent variable in place of u), we find

$$(32) \qquad F = f_2 \left(u, v_1, \frac{u - u_1}{v - v_1} \right) + f_3 \left(v, u_1, \frac{v - v_1}{u - u_1} \right).$$

Here f_2 and f_3 are arbitrary, except of course that their sum must be real when we now restrict x, y, x_1, y_1 to real values. Since the three variables appearing in f_3 now become the conjugates of those in f_2, an equivalent form of (32) is

$$(33) \qquad F = \Re \left[f \left(u, v_1, \frac{u - u_1}{v - v_1} \right) \right],$$

where f is arbitrary (this follows from the fact that the real part of an analytic function of complex variables is also the real part of some analytic function of the conjugates of those variables). The result given in Theorem 5 is identical with (33).

In case VIII, $F_{uv} = F_{u_1 v_1} = 0$ by (27). This means that

$$(34) \qquad F = f_1(u, u_1) + f_2(u, v_1) + f_3(v, u_1) + f_4(v, v_1),$$

where the f's are arbitrary. Substituting (34) into the third equation in (27) and noting the separation of variables, we have

$$(35) \qquad (u - u_1)^2 (f_1)_{uu_1} = -(v - v_1)^2 (f_4)_{vv_1} = \lambda,$$

where λ is a complex constant. Integrating (35) and substituting in (34), we find

$$(36) \qquad F = f_4(u, v_1) + f_5(v, u_1) + \lambda \log \frac{u - u_1}{v - v_1}.$$

The functions f_4 and f_5 and the constant λ in (36) are arbitrary, subject to the condition that when we now make x, y, x_1, y_1 real, F also becomes real. Since the variables in f_5 are now the conjugates of those in f_4, and since the coefficient of λ is now a pure imaginary, it follows that a form of the answer equivalent to (36) is

$$(37) \qquad F = \Re[f(u, v_1)] + \Re \left[\lambda \log \frac{u - u_1}{v - v_1} \right],$$

where the function f and the constant λ are completely arbitrary. The results stated in case VIII of Theorem 5 are readily derived from (37). This concludes the proof of Theorem 5.

Mixed Groups.

5. The previous theorems deal either with groups and sets of conformal substitutions, or else with sets of reverse conformal substitutions. There remains the case of mixed groups or sets, consisting of both direct and reverse conformal substitutions. In particular, we may ask: what functions are rendered harmonic by all the substitutions of the total (direct and reverse) conformal group? By Theorem 2 and its analogue for reverse conformal substitutions, the required functions are those which are at once biharmonic and anti-biharmonic; that is, they are those functions which satisfy equations (2) and also (19). Hence we must have

$$(38) \qquad F_{xx} + F_{yy} = F_{xx_1} = F_{xy_1} = F_{yx_1} = F_{yy_1} = F_{x_1x_1} + F_{y_1y_1} = 0.$$

The solution of (38) is

$$(39) \qquad F = f(x, y) + g(x_1, y_1),$$

where f and g are arbitrary harmonic functions of two variables. This proves

Theorem 6. *The only functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every substitution of the total (direct and reverse) conformal group, are those which are equal to the sum of a harmonic function of x and y and a harmonic function of x_1 and y_1.*

Finally, we take up the question of determining what functions are made harmonic by all the substitutions of a subgroup of the total conformal group, in the case of these two mixed subgroups:

(IX) The total group of $2\infty^3$ rigid motions, generated by the set VII of line symmetries (the group of ordinary plane geometry).

(X) The inversion group, generated by the set VIII of transformations by reciprocal radii vectores. This consists of the $2\infty^6$ linear fractional transformations of z and linear fractional transformations of $\bar{z}$.

A function F which is made harmonic by the total group of motions IX must satisfy the equations deduced in case III for direct motions, together with the corresponding equations for reflected motions. The latter equations are derived from the former by replacing P_2, P_3 by P'_2, P'_3, as we see by comparing (22) with (8). Hence F is determined by the system

$$(40) \qquad F_{xx} + F_{yy} + F_{x_1x_1} + F_{y_1y_1} = F_{xx_1} = F_{xy_1} = F_{yx_1} = F_{yy_1} = 0.$$

The last four equations mean that F is separable in the pairs x, y and x_1, y_1. Then using this separation of variables in the first equation, we see as in the proof of III that F equals the sum of a harmonic function of x and y,

a harmonic function of x_1 and y_1, and a function of the special form $c(x^2+y^2-x_1^2-y_1^2)$.

In case X, it is sufficient to recall that the inversion group contains all similitudes, both direct and reverse. Then by Theorem 2′ and its analogue for reverse similitudes, F must be both biharmonic and anti-biharmonic. Hence the answer for the inversion group is the same as for the total conformal group. We have now proved

THEOREM 7. *The class of functions $F(x, y, x_1, y_1)$ which are converted into harmonic functions of x and y by every substitution $x_1 = \phi(x, y)$, $y_1 = \psi(x, y)$ of a mixed group of (direct and reverse) conformal substitutions is, in the case of the groups* IX *and* X, *the following:*

(IX) *(Total group of motions):*
$F = f(x, y) + g(x_1, y_1) + c(x^2+y^2-x_1^2-y_1^2)$, *where f and g are arbitrary harmonic functions and c is an arbitrary constant.*

(X) *(The inversion group): $F = f(x, y) + g(x_1, y_1)$, where f and g are arbitrary harmonic functions.*

Since the group of motions IX is generated by the set of line symmetries VII, and therefore includes this set, the class of functions F which are made harmonic by the group IX must be included in the class of functions made harmonic by the set VII. The same thing may be said for the inversion group X and the set of inversions VIII. That is, the answers in Theorem 5 should include, respectively, the answers in Theorem 7. For VIII and X this is evidently so, for the answer to case X can also be written in the form

$$F = \Re[f_1(z) + f_2(\bar{z}_1)], \tag{41}$$

and this is of the type VIII in Theorem 5. For cases VII and IX the fact is less obvious, but it follows from the relation

$$x^2+y^2-x_1^2-y_1^2 = \Re\left[z^2\frac{\bar{z}-\bar{z}_1}{z-z_1} - \bar{z}_1^2\frac{z-z_1}{\bar{z}-\bar{z}_1}\right].[8] \tag{42}$$

QUADRATIC BIHARMONIC FUNCTIONS.

6. By Theorem 2, not every (direct) conformal substitution can reduce a given function $F(x, y, x_1, y_1)$ to a harmonic function of x and y (unless of course F happens to be biharmonic). But for a *given* non-biharmonic F, there may still be some subset of the conformal group which makes F harmonic.

[8] This relation can be derived by following the steps of the proof of VII, taking F to be $x^2+y^2-x_1^2-y_1^2$. Of course it can also be verified directly.

We determine this subset in the case that F is a quadratic polynomial. There is no need for the subset to constitute a group, and we find that as a rule groups do not appear. We first establish the following

LEMMA. *If F is a quadratic polynomial in four variables, and if a conformal substitution T reduces F to a harmonic function, then T must be linear (unless F is biharmonic).*

For (8) must hold, with the P_iF reduced to constants since F is a quadratic polynomial. Differentiating (8) with respect to x and y, and rearranging terms, we have:

$$\begin{aligned}\phi_{xx}(\phi_x P_4 + P_2)F + \phi_{xy}(\phi_y P_4 - P_3)F = 0,\\ \phi_{xy}(\phi_x P_4 + P_2)F + \phi_{yy}(\phi_y P_4 - P_3)F = 0.\end{aligned} \tag{43}$$

Now since T is conformal, $\phi_{xx} + \phi_{yy} = 0$. If the determinant $\phi_{xx}\phi_{yy} - \phi_{xy}^2$ of (43) vanishes, then using $\phi_{yy} = -\phi_{xx}$, we have $\phi_{xx}^2 + \phi_{xy}^2 = 0$, so that $\phi_{xx} = \phi_{xy} = 0$. Then ϕ_{yy} is also zero, and T must be linear. If the determinant of the homogeneous system (43) does not vanish, then

$$\phi_x P_4 F + P_2 F = \phi_y P_4 F - P_3 F = 0.$$

This means that ϕ_x and ϕ_y are constants, so that ϕ, and therefore ψ, is linear; unless $P_4F = P_2F = P_3F = 0$. In the latter case we have from (8), $P_1F = 0$; hence F is biharmonic. This proves the lemma.

Now if a linear conformal substitution (9) makes F harmonic, the condition which the coefficients a, b in (9) must satisfy is, as in (10),

$$(a^2 + b^2)P_4F + 2aP_2F + 2bP_3F + P_1F = 0, \tag{44}$$

where $P_1F, \cdots, P_4F$ are independent constants, not all zero unless F is biharmonic. If we interpret a, b as the cartesian coördinates of a point in the plane, (44) is the general circle [9] equation; its (real) graph is either a circle, or a straight line, or a point, or does not exist in the real domain. Corresponding to each point (a, b) of the graph there are ∞^2 similitudes (9), since c, d are unrestricted. We can state this result in the following way:

THEOREM 8. *If F is a quadratic polynomial in x, y, x_1, y_1, the set of (direct) conformal substitutions which make F harmonic is one of four kinds, depending on the coefficients of F:*

---

[9] This may be called the *auxiliary circle* and we shall use it elsewhere. The proof of Theorem 2 can easily be changed to yield this stronger form of Theorem 2′: *If F is made harmonic by $4\infty^2$ similitudes situated "at four points" (a, b), and if these four points are not co-circular, then F is biharmonic.* For any given function $F(x, y, x_1, y_1)$ and any given bipoint (x, y), (x_1, y_1) we have a definite auxiliary circle.

(A) *The entire conformal group (namely, when F is biharmonic).*
(B) ∞^3 *similitudes " on a circle " or " on a straight line."*
(C) ∞^2 *similitudes " at a point."*
(D) *None.*

We remark that (B) and (D) constitute the general case; (C) is special, and (A) is still more special. For (C) arises only when the constants P_iF satisfy a certain relation, expressing the fact that the radius of the circle (44) is zero; and (A) requires the four conditions $P_iF = 0$.

The only groups under (B) are the group of rigid motions III (" on the circle " $a^2 + b^2 = 1$), and the group of magnifications II (" on the line " $b = 0$, excluding the origin); under (C) there is only the group of translations IV (" at the point " $(1, 0)$). To prove this we use the fact that if a_1, b_1 and a_2, b_2 are the coefficients of two similitudes (9), and if a_3, b_3 are the coefficients of the product of those two similitudes, then

$$a_3 + ib_3 = (a_1 + ib_1)(a_2 + ib_2).$$

This is evident from the complex form of (9), namely:

$$z_1 = (a + ib)z + (c + id).$$

Hence if the ∞^3 similitudes corresponding to points (a, b) on a circle are to form a group, the set of complex numbers $a + ib$ must form a group with respect to the operation of multiplication. Now it is easy to see that the successive powers of a complex number cannot lie on any circle, unless the modulus of the number is unity (or zero). Therefore the unit circle $a^2 + b^2 = 1$ is the only one possible; and this circle does in fact correspond to the group of motions. Similarly, the only straight line group is the real axis $b = 0$ (excluding 0); and the only point group (other than zero) is the point unity. These correspond, respectively, to the magnifications $z_1 = az + (c + id)$, and the translations $z_1 = z + (c + id)$.

Extensions.

7. We remark in conclusion that most of our theorems can be extended to *multi-harmonic* functions connected with functions of n (instead of two) complex variables. There is also a somewhat analogous theory of the reduction of polygenic (non-analytic) functions of complex variables to monogenic functions by monogenic substitutions.

Columbia University,
New York.

THE PROBLEM OF BOLZA IN THE CALCULUS OF VARIATIONS IN PARAMETRIC FORM.[1]

By Magnus R. Hestenes.[2]

1. Introduction. In a Riemannian space $\mathfrak{R}$ as defined by Morse (X, pp. 107-108)[3] with local coördinates $(x) = (x^0, x^1, \cdots, x^n)$ the problem of Bolza can be defined as follows. Let $(x^1) = (x^{01}, x^{11}, \cdots, x^{n1})$ and $(x^2) = (x^{02}, x^{12}, \cdots, x^{n2})$ be respectively the end points 1 and 2 of the arc

$$x^i = x^i(t) \qquad (t^1 \leqq t \leqq t^2;\ i = 0, 1, \cdots, n). \tag{1.1}$$

We seek to find in a class of arcs (1.1) and sets of constants $(\alpha) = (\alpha^1, \cdots, \alpha^r)$ satisfying the differential equations and end conditions

$$\phi_\beta(x, \dot{x}) = 0 \qquad (\beta = 1, \cdots, m < n), \tag{1.2}$$

$$x^{is} = x^{is}(\alpha^1, \cdots, \alpha^r) \qquad (s = 1, 2) \tag{1.3}$$

one which minimizes a functional of the form

$$J = \theta(\alpha^1, \cdots, \alpha^r) + \int_{t^1}^{t^2} f(x, \dot{x})\, dt.$$

The problem of Bolza in non-parametric form has been studied by a number of writers.[4] The only treatments of the parametric problem known to the author are those of Eshleman (III) and Hefner (VII) in the problem of Lagrange.

Sufficient conditions for a minimum in the non-parametric problem without normality assumptions on the subintervals were first obtained by the author while he was a Research Assistant to Professor Bliss. These results were later extended (XIII) to the case in which no normality assumptions were made. This extension is a trivial one. More recently, Morse (XIV) and Reid (XV) obtained independently further sufficient conditions involving conjugate points which are equivalent to those given by the author (See XVI). In the present paper we give similar sufficiency conditions for a minimum in the parametric case.

---

[1] Presented to the American Mathematical Society, September 6, 1934.

[2] National Research Fellow 1933-4.

[3] Roman numerals in parentheses refer to the list of references at the end of this paper.

[4] For references other than those given at the end of this paper see (XIII).

2. Preliminary remarks. In the present paper we shall assume that two systems of coördinates (x) and (z) which admissibly represent the same neighborhood of a point of our Riemannian space $\mathcal{R}$ are connected by an equation of the form

$$(2.1) \qquad z^i = z^i(x^0, x^1, \cdots, x^n) \qquad (i = 0, 1, \cdots, n)$$

in which the functions $z^i(x)$ are of class[5] C^4 and possess a non-vanishing jacobian. Let $(r) = (r^0, r^1, \cdots, r^n)$ be a contravariant vector, that is, suppose that under the transformation (2.1) the vector (r) is transformed into a vector (σ) according to the law[6]

$$(2.2) \qquad \sigma^i = \frac{\partial y^i}{\partial x^j} r^j. \qquad (i, j = 0, 1, \cdots, n).$$

The functions $f(x, r)$, $\phi_\beta(x, r)$ appearing in equations (1.2) and (1.4) are assumed to be defined over $\mathcal{R}$ for all sets $(r) \neq (0)$ and to be invariants under the transformations (2.1) and (2.2). Moreover these functions are assumed to be of class C^3 and to satisfy the homogeneity relations

$$(2.3) \quad f(x, kr) = kf(x, r), \qquad \phi_\beta(x, kr) = k\phi_\beta(x, r) \qquad (k > 0).$$

By a *regular arc* is meant a set of points in $\mathcal{R}$ defined locally by equations of the form (1.1) of class C^1 and having $(\dot{x}) \neq (0)$. By a *differentially admisssible arc* is meant a continuous succession of a finite number of regular arcs satisfying the differential equations (1.2). By an *admissible arc* is meant a differentiably admissible arc (1.1) and a set of constants $(\alpha) = (\alpha^1, \alpha^2, \cdots, \alpha^r)$ which satisfy the end conditions (1.3).

We are concerned with a particular admissible arc g without corners. As was noted by Morse we may assume without loss of generality that the arc g does not intersect itself. We suppose further that the matrix $\| \phi_{\beta r^i} \|$ has rank m along g and that the set (α) belonging to g is the set $(\alpha) = (0)$. The functions $x^{i1}(\alpha)$, $x^{i2}(\alpha)$, $\theta(\alpha)$ appearing in equations (1.3) and (1.4) are assumed to be of class C^2 near $(\alpha) = (0)$.

Our problem is then that of finding under what conditions the arc g will furnish a minimum to the functional J relative to neighboring admissible arcs. These conditions must be independent of the admissible coördinates (x) and of admissible parameters t and (α) used. Moreover, it can be shown that a neighborhood of g can be covered by a single coördinate system. In fact this

---

[5] A function $f(x^0, x^1, \cdots, x^n)$ is said to be of class C^n if it possesses continuous partial derivatives of the first n orders.

[6] The tensor analysis summation convention is used.

coördinate system can be chosen so that $\dot{x}^0 > 0$ along g. In our proofs we shall suppose that such a coördinate system has been chosen. Our results, however, are valid for all coördinate systems.

3. First necessary conditions. If we consider our problem as a non-parametric problem in tx-space we obtain at once the results of this section from the non-parametric case (XIII).

THEOREM 3.1. *If g is a minimizing arc then there exists a function $F = \lambda^0 f + \lambda^\beta(t)\phi_\beta$ such that the Euler-Lagrange equations*

$$P_i = (d/dt)F_{r^i} - F_{x^i} = 0, \qquad \phi_\beta = 0 \qquad (i = 0, 1, \cdots, n) \tag{3.1}$$

hold at each point of g. Moreover on g the equation [7]

$$[F_{r^i}dx^i]_1{}^2 + \lambda^0 d\theta = 0 \tag{3.2}$$

is an identity in $d\alpha_h$ when the differentials dx^{i1}, dx^{i2}, $d\theta$ are expressed in terms of $d\alpha_h$. The multiplier λ^0 is a constant. The multipliers $\lambda^\beta(t)$ are continuous and the elements of the set λ^0, $\lambda^\beta(t)$ do not vanish simultaneously at any point of g.

It is readily proved that under the transformation of admissible coördinates (x) the functions P_i in equations (3.1) are covariant components of a vector and that the first member of the identity (3.2) is an invariant. It follows that the multipliers λ^0, $\lambda^\beta(t)$ belonging to g are invariants. Moreover the number of linearly independent multipliers of the form $\lambda^0 = 0$, $\lambda^\beta(t)$ with which a subarc g_1 of g satisfies the equations (3.1) is clearly a numerical invariant and is called the *order of anormality of g on g_1 relative to the equations* (3.1). Similarly the number of linearly independent multipliers of the form $\lambda^0 = 0$, $\lambda^\beta(t)$ with which g satisfies the conditions (3.1) and (3.2) is also a numerical invariant and is called the *order of anormality of g relative to the equations* (3.1) *and* (3.2). If there are no multipliers of the latter type g is said to be normal and the multiplier λ^0 in Theorem 3.1 can be taken to be unity and in this form the multipliers are unique.

We come now to the *necessary condition of Weierstrass* (see Graves, VIII).

THEOREM 3.2. *If g is a normal minimizing arc then at each element $(x, \dot{x}, \lambda)$ on g the inequality*

$$E(x, \dot{x}, \lambda, r) \geqq 0$$

[7] Here and elsewhere the symbol $[\ ]_1{}^2$ denotes the value of $[\ ]$ at the point 2 minus the value of $[\ ]$ at the point 1.

must hold for every set $(x, r) \neq (x, k\dot{x})$ $(k > 0)$ *which satisfies the equations* $\phi_\beta = 0$ *and gives the matrix* $\| \phi_{\beta r^i} \|$ *rank* m, *where*

$$E(x, \dot{x}, \lambda, r) = F(x, r, \lambda) - r^i F_{r^i}(x, \dot{x}, \lambda). \tag{3.3}$$

The *necessary condition of Clebsch* is described in the following theorem:

THEOREM 3.3. *If* g *is a normal minimizing arc then at each element* $(x, \dot{x}, \lambda)$ *on* g *the inequality*

$$F_{r^i r^j} \pi^i \pi^j \geqq 0 \qquad (i, j = 0, 1, \cdots, n) \tag{3.4}$$

must hold for all sets $(\pi) \neq (\tau \dot{x})$ *which satisfy the equations*

$$\phi_{\beta r^i} \pi^i = 0. \tag{3.5}$$

The Weierstrass E-function (3.3) is clearly an invariant. So also are the first members of the relations (3.4) and (3.5) provided that (π) is a contravariant vector.

We shall now define a new analogue of the Weierstrass function F_1. In order to do so we first note that the determinant

$$\begin{vmatrix} F_{r^i r^j} & \phi_{\beta r^i} \\ \phi_{\beta r^j} & 0 \end{vmatrix} \tag{3.6}$$

is identically zero along g, as follows readily from the equations

$$F_{r^i r^j} r^j = 0, \qquad \phi_\beta = r^j \phi_{\beta r^j} \tag{3.7}$$

which are found by differentiating the equations (2.3) for k and r^i and setting $k = 1$. Let $b_p{}^i(x, r)$ $(p = 1, \cdots, n)$ be a set of n contravariant vectors homogeneous of order zero in (r) and having its determinant $| b_p{}^i r^i |$ different from zero. By the *Weierstrassian function* F_1 is meant the determinant

$$F_1(x, r, \lambda) = \begin{vmatrix} b_p{}^i F_{r^i r^j} b_q{}^j & \phi_{\beta r^i} b_p{}^i \\ \phi_{\beta r^j} b_q{}^j & 0 \end{vmatrix}. \tag{3.8}$$

One readily verifies that the function F_1 is different from zero along g if and only if the matrix (3.6) has rank $n + m$ on g. It can be shown further that the function F_1 here defined differs from the function F_1 used hitherto by a non-vanishing factor (cf. VII, p. 104; X, pp. 112-113). Moreover the function F_1 here defined is an absolute invariant whereas the function F_1 used hitherto is a relative invariant of the second order.

By an *extremal arc* is meant a differentially admissible arc and a set of multipliers, not vanishing simultaneously

$$x^i = x^i(t), \quad \lambda^0, \quad \lambda^\beta = \lambda^\beta(t) \qquad (t^1 \leqq t \leqq t^2)$$

possessing continuous derivatives $\dot{x}^i$, $\ddot{x}^i$, $\dot{\lambda}^\beta$ and satisfying the equations (3.1). An extremal is said to be *non-singular* if the function F_1 is different from zero at each point on it.

If the arc g under consideration satisfies the equations (3.1) with a set of multipliers λ^0, $\lambda^\beta(t)$ and is such that $F_1(x, \dot{x}, \lambda)$ is different from zero on it, then g is an extremal arc. Moreover the parameter t can be chosen so that the functions $x^i(t)$ defining g are of class C^3 and the functions $\lambda^\beta(t)$ of class C^2 (cf. VII, p. 103).

In the sequel we shall assume that g is an extremal arc having $\lambda^0 = 1$ and satisfying the transversality condition (3.2), *unless otherwise expressly stated.* It will be understood further that the parameter t has been chosen so as to give the maximum number of derivatives.

4. The second variation and the accessory minimum problem. By considering our problem as a non-parametric problem in tx-space with t^1 and t^2 fixed it is found that along g the second variation of the functional J takes the form

$$J_2(\eta, w) = b_{hl} w^h w^l + \int_{t^1}^{t^2} 2\omega(\eta, \dot{\eta})\, dt, \tag{4.1}$$

where

$$2\omega = F_{x^i x^j} \eta^i \eta^j + 2F_{x^i r^j} \eta^i \dot{\eta}^j + F_{r^i r^j} \dot{\eta}^i \dot{\eta}^j,$$
$$b_{hl} = \theta_{hl} + [F_{r^i} x^{is}_{hl}]_1{}^2,$$

the subscripts h, l in the second member of the last equation denoting differentiation with respect to α_h, α_l and setting $(\alpha) = (0)$. The variations $\eta^i(t)$ are assumed to be continuous and to have continuous derivatives except possible at a finite number of values of t on $t^1 t^2$ and to satisfy with the constants w^h the equations

$$\Phi_\beta(\eta) = \phi_{\beta x^i} \eta^i + \phi_{\beta r^i} \dot{\eta}^i = 0, \tag{4.2}$$

$$\eta^{is} - x_h{}^{is} w^h = 0 \qquad (s = 1, 2), \tag{4.3}$$

where the coefficients of η^i, $\dot{\eta}^i$ in (4.2) are evaluated along g and the subscript h denotes differentiation with respect to α_h at $(\alpha) = (0)$. Such variations (η, w) will be termed *admissible variations.* As was seen by Morse (X, p. 123) the variations η^i, $x_h{}^{i1}$, $x_h{}^{i2}$ define respectively contravariant vectors. It follows that the condition (4.3) is independent of the coördinate system

(x) used. Similarly the functions $\Phi_\beta(\eta)$ are invariants and the functional $J_2(\eta, w)$ is an invariant subject to the condition (4.3), as can be seen from the way in which they were derived. The functions 2ω and b_{hl} however are not invariants, as one readily verifies.

The following necessary condition follows at once from the non-parametric case (XIII, p. 797):

THEOREM 4.1. *If g is a normal minimizing extremal arc the second variation (4.1) of J must satisfy along g the condition $J_2(\eta, w) \geqq 0$ for every set of admissible variations (η, w) having continuous second derivatives except possibly at a finite number of points on t^1t^2.*

The theorem just described suggests the study of the problem of minimizing the functional $J_2(\eta, w)$ in the class of admissible variations η^i, w^h. The problem is a non-parametric problem of Bolza and will be termed *the accessory minimum problem.* The analogues of equations (3.1) and (3.2) are the following:

$$L_i(\eta, \mu) \equiv (d/dt)\Omega_{\dot\eta^i} - \Omega_{\eta^i} = 0, \qquad \Phi_\beta = 0, \tag{4.4}$$

$$\zeta_i{}^2 x_h{}^{i2} - \zeta_i{}^1 x_h{}^{i1} + b_{hl}w^l = 0 \qquad (h, l = 1, \cdots, r), \tag{4.5}$$

where

$$\Omega = \omega + \mu^\beta\Phi_\beta, \qquad \zeta_i = \Omega_{\dot\eta^i}.$$

Here we have taken the constant μ^0 analogous to λ^0 in (3.1) to be unity, which is permissible since the accessory minimum problem can always be made normal. The equations (4.4) are known as the *accessory differential equations,* the equations (4.5) as the *secondary transversality conditions.* The extremals with $\mu^0 = 1$ for this problem will be called *secondary extremals.*

By an argument like that given by Morse (X, p. 123) it is readily seen that on g the functions $L_i(\eta, \mu)$ are covariant components of a vector and that the first members of equations (4.5) are invariants subject to the conditions (4.3). The multipliers μ^β belonging to a secondary extremal η^i, μ^β are therefore invariants, the functions η^i being contravariant as was seen above.

If $\tau(t)$ is a continuous function possessing a continuous derivative except possibly at a finite number of points on t^1t^2, then as is well known (VII, p. 114) the functions

$$\eta^i = \tau\dot{x}^i, \qquad \mu^\beta = \tau\lambda^\beta \tag{4.6}$$

satisfy equations (4.4). It follows that if η^i, μ^β is a secondary extremal so also is $\eta^i - \tau\dot{x}^i$, $\mu^\beta - \tau\dot\lambda^\beta$ provided the function $\tau(t)$ is of class C^2. It

should be noted also that the equations (4. 4) are not all independent since the equation

$$\dot{x}^i L_i(\eta, \mu) = \lambda^\beta \Phi_\beta(\eta) \tag{4. 7}$$

is an identity in the variables η^i, $\dot{\eta}^i$, $\ddot{\eta}^i$, μ^β, $\dot{\mu}^\beta$, as one readily verifies with the help of known consequences of the homogeniety conditions (2. 3) such as equations (3. 7).

The following formula will be useful in the next section:

$$J_2(\eta - \tau\dot{x}, w) = J_2(\eta, w) + [\tau^2 \dot{x}^i F_{x^i} + 2\tau F_{x^i}\eta_i]_1{}^2. \tag{4. 8}$$

This relation holds for all functions η^i, w^h, τ having the continuity properties described in the above paragraphs and satisfying equations (4. 2). In order to establish this relation we note that

$$\begin{aligned} 2\omega(\eta - \tau\dot{x}) &= 2\Omega(\eta - \tau\dot{x}) \\ &= 2\omega(\eta) + 2[\eta^i \Omega_{\eta^i}(\tau\dot{x}) + \dot{\eta}^i \Omega_{\dot{\eta}^i}(\tau\dot{x})] + 2\Omega(\tau\dot{x}) \end{aligned} \tag{4. 9}$$

the multipliers μ^β in Ω being $\tau\lambda^\beta$. Noting that by virtue of equations (3. 7) and (3. 1)

$$\Omega_{\dot{\eta}^i}(\tau\dot{x}) = \tau(d/dt)F_{\dot{x}^i} = \tau F_{x^i} \tag{4. 10}$$

and recalling that the functions (4. 6) satisfy equations (4. 4) it is found that

$$\eta^i \Omega_{\eta^i}(\tau\dot{x}) + \dot{\eta}^i \Omega_{\dot{\eta}^i}(\tau\dot{x}) = (d/dt)[\eta^i \Omega_{\dot{\eta}^i}(\tau\dot{x})] = (d/dt)(\tau F_{x^i}\eta^i). \tag{4. 11}$$

Similarly from the well known identity

$$2\Omega = \eta^i \Omega_{\eta^i} + \dot{\eta}^i \Omega_{\dot{\eta}^i} + \mu^\beta \Omega_{\mu^\beta}$$

we find that

$$2\Omega(\tau\dot{x}) = (d/dt)(\tau^2 F_{x^i}\dot{x}^i). \tag{4. 12}$$

The relation (4. 8) now follows at once from the equation (4. 9), (4. 11), and (4. 12).

5. The second variation in normal form. Since the determinant (3. 6) is always zero it is clear that the accessory minimum problem described in the last section is singular. In the present section we shall set up an equivalent accessory minimum problem which is always non-singular whenever the function F_1 is different from zero along g. The method used is essentially a generalization of a method due to Weierstrass (See I, pp. 224-226) and seems to be particularly adaptable to the present paper (cf. II, IV, VII, XII).

Let $a_i{}^p(x)$ $(p = 1, \cdots, n)$ be a set of n linearly independent covariant vectors of class C^2 near g and satisfying the equations $a_i{}^p\dot{x}^i = 0$ along g. For

example we may choose these vectors so as to form on g a set of n mutually orthogonal vectors normal to g, as does Weierstrass in the two dimensional case. It is clear that the functions $u^p(t)$ defined along g by the equations

$$u^p = a_i^p \eta^i \qquad (p = 1, \cdots, n) \tag{5.1}$$

are invariants relative to transformations of admissible coördinates (x). Moreover the variations η^i and $\eta^i - \tau \dot{x}^i$ determine the same functions u^p. If now we select a covariant vector $a_i(x)$ such that along g we have $a_i \dot{x}^i \neq 0$, then there exist a set of n linearly independent contravariant vectors $b_p^i(x)$ of class C^2 with $a_i b_p^i = 0$ along g and such that the equations

$$\eta^i = b_p^i u^p + \tau \dot{x}^i, \qquad \tau = a_i \eta^i / a_j \dot{x}^j \tag{5.2}$$

hold along g, whenever the equations (5.1) are satisfied. It follows that a set of admissible variations $\eta^i(t)$, w^h determine a unique set of functions $u^p(t)$, w^h, $\tau(t)$ of class D^1 satisfying the equations

$$\Phi_\beta{}^0(u) = \Phi_\beta(b_p u^p + \tau \dot{x}) = \Phi_\beta(b_p u^p) = 0, \tag{5.3}$$

$$u^{ps} - c_h{}^{ps} w^h = 0, \qquad \tau^s = d_h{}^s w^h \tag{5.4}$$

where (s not summed)

$$\begin{aligned} c_h{}^{ps} &= a_i{}^{ps} x_h{}^{is}, \qquad d_h{}^s = a_i{}^s x_h{}^{is} / a_j{}^s \dot{x}^{js} \\ x_h{}^{is} &= b_p{}^{is} c_h{}^{ps} + d_h{}^s \dot{x}^{is} \end{aligned} \qquad (s = 1, 2). \tag{5.5}$$

Conversely every set of functions $u^p(t)$, w^h, $\tau(t)$ of class D^1 satisfying equations (5.3) and (5.4) determines by means of equations (5.2) a set of admissible variations $\eta^i(t)$, w^h. A set of functions $u^p(t)$, w^h of the type just described will be said to be a set of *normal admissible variations.*

Under the transformation (5.2) the second variation takes the form

$$J_2{}^0(u, w) = b_{hl}{}^0 w^h w^l + \int_{t^1}^{t^2} 2\omega^0(u, \dot{u})\, dt, \tag{5.6}$$

where

$$\begin{aligned} 2\omega^0 &= 2\omega(b_p u^p), \\ b_{hl}{}^0 &= b_{hl} + [\dot{x}^i F_{x^i} d_h{}^s d_l{}^s + F_{x^i}(x_h{}^{is} d_l{}^s + x_l{}^{is} d_h{}^s)]_1{}^2, \end{aligned} \tag{5.7}$$

as is easily seen with the help of the formula (4.8). The expression (5.6) will be termed the *normal form of the second variation.* It is invariant subject to the conditions (5.4), although the functions $b_{hl}{}^0$, $2\omega^0$ are not invariants.

The following result is immediate:

LEMMA 5.1. *The functional $J_2(\eta, w)$ is positive (non-negative) for all admissible variations $(\eta, w) \neq (\tau\dot{x}, 0)$ if and only if the functional $J_2{}^0(u, w)$ is positive (non-negative) for all normal admissible variations $(u, w) \neq (0, 0)$.*

In view of this last lemma it is clear that the accessory minimum problem is equivalent to that of minimizing the functional $J_2{}^0(u, w)$ in the class of normal variations (u, w). This new problem will be termed the *normal accessory minimum problem.* It is of the type already studied by the author (XIII, p. 798) and its properties are known. It remains to interpret these properties in terms of the original problem. The analogues of equations (4.4) and (4.5) are the following

$$(5.8) \qquad L_p{}^0(u, \rho) = (d/dt)\Omega_{\dot{u}^p}{}^0 - \Omega_{u^p}{}^0 = 0, \qquad \Phi_\beta{}^0 = 0$$

$$(5.9) \qquad v_p{}^2 c_h{}^{p2} - v_p{}^1 c_h{}^{p1} + b_{hl}{}^0 w^l = 0,$$

where

$$\Omega^0 = \omega^0 + \rho^\beta \Phi_\beta{}^0, \qquad v_p = \Omega_{\dot{u}^p}{}^0.$$

An extremal (u, ρ) for the problem just described will be termed a *normal secondary extremal.*

LEMMA 5.2. *Every secondary extremal (η, μ) determines a unique normal secondary extremal (u, ρ) and a function $\tau(t)$ of class C^2 such that the equations* (5.1) *and*

$$(5.10) \qquad \eta^i = b_p{}^i u^p + \tau\dot{x}^i, \qquad \mu^\beta = \rho^\beta + \tau\dot{\lambda}^\beta, \qquad \tau = a_i\eta^i / a_j\dot{x}^j$$

hold along g, and conversely. The corresponding values of ζ_i, v_p are connected by the formulas

$$(5.11) \qquad \zeta_i = a_i{}^p v_p + \tau F_{x'^i}, \qquad v_p = b_p{}^i(\zeta_i - \tau F_{x'^i}).$$

This result will follow at once if we note that under the transformation (5.10) the functions L_i, $L_p{}^0$ of equations (4.4) and (5.8) are connected by the formulas

$$(5.12) \qquad L_p{}^0 = b_p{}^i L_i, \qquad L_i = a_i{}^p L_p{}^0.$$

This result is readily established by substitution.

In view of equations (5.12) it is clear that the functions $L_p{}^0(u, \rho)$ are invariants. Similarly the first members of equations (5.9) are invariants subject to the conditions (5.4) since under the transformation (5.10) they are equal to the first members of equations (4.5), as one readily verifies with the help of equations (5.5), (5.7), and (5.11).

Suppose now that g is non-singular. The determinant (3.8) is then different from zero along g and the normal secondary extremals (u, ρ) are therefore also non-singular. The following result is an immediate consequence of Lemma 5.2 and known theorems concerning the solutions of equations (5.8).

THEOREM 5.1. *If the extremal g is non-singular then the rank of the matrix*

$$\left\| \begin{matrix} \eta_j{}^i & \dot{x}^i \\ \zeta_{ij} & F_{x'^i} \end{matrix} \right\|, \qquad \zeta_{ij} = \Omega_{\dot{\eta}^i}(\eta_j, \dot{\eta}_j, \mu_j) \tag{5.13}$$

formed for p secondary extremals $\eta_j{}^i$, $\mu_j{}^\beta$ $(j = 1, \cdots, p)$ is the same at each point on t^1t^2. Moreover every secondary extremal (η, μ) is expressible linearly in the form

$$\eta^i = \eta_j{}^i a^j + \tau \dot{x}^i, \qquad \mu^\beta = \mu_j{}^\beta a^j + \tau \lambda^\beta \tag{5.14}$$

in terms of $p = 2n$ secondary extremals $\eta_j{}^i$, $\mu_j{}^\beta$ whose matrix (5.13) has rank $2n + 1$, the a's being constants and the function $\tau(t)$ being of class C^2. The corresponding value of ζ_i is

$$\zeta_i = \zeta_{ij} a^j + \tau F_{x'^i}. \tag{5.15}$$

With the help of the last theorem and Lemma 5.2 one can readily establish analogues of necessary conditions for the second variation to be non-negative as given by the author for the non-parametric case (XIII, pp. 801-803).

We come now to the notion of conjugate points. A value $t^3 \neq t^1$ will be said to define a point 3 *conjugate* to 1 on g if there exists a secondary extremal (η, μ) having $\eta^i(t^1) = \eta^i(t^3) = 0$ and $(\eta) \not\equiv (\tau \dot{x})$ on t^1t^3. By virtue of the equations (5.10) this is equivalent to the condition that there exists a normal secondary extremal (u, ρ) with $u^p(t^1) = u^p(t^3) = 0$ and $(u) \not\equiv (0)$ on t^1t^3. The following theorem is an immediate consequence of equations (5.10) and a theorem of Hestenes (XIII, p. 804).

THEOREM 5.2. *Suppose the extremal g is non-singular. If $\eta_j{}^i$, $\mu_j{}^\beta$ $(j = 1, \cdots, 2n)$ form $2n$ secondary extremals whose matrix (5.13) has rank $2n + 1$ then a value $t^3 \neq t^1$ defines a point 3 conjugate to the point 1 on g if and only if the matrix*

$$\left\| \begin{matrix} \eta_j{}^i(t^3) & \dot{x}^i(t^3) & 0 \\ \eta_j{}^i(t^1) & 0 & \dot{x}^i(t^1) \end{matrix} \right\|$$

has rank $r < 2n - q + 2$, where q is the order of anormality of g in its subarc determined by t^1t^3.

6. The accessory boundary value problem. Let $K_{ij}(x, r)$ be a set of

functions of class C^1 in a neighborhood of the values $(x, r) = (x, \dot{x})$ on g, and such that along g we have $K_{ij} = K_{ji}$ and

$$(6.1) \qquad K_{ij}\eta^i\eta^j > 0 \qquad (\eta) \neq (\tau\dot{x}), \qquad K_{ij}r^j = 0.$$

Moreover these functions are assumed to be positively homogeneous of the first order in (r) and to be the components of a covariant tensor of the second order. For example, we can in general select $K_{ij} = F_{r^ir^j}$. By the accessory boundary value problem is meant the system of equations

$$(6.2) \qquad \begin{aligned} &M_i(\eta, \mu, \sigma) = L_i(\eta, \mu) + \sigma K_{ij}\eta^j = 0, && \Phi_\beta = 0, \\ &\eta^{is} - x_h^{is}w^h = 0 && (i, j = 0, 1, \cdots, n;\ s = 1, 2) \\ &\zeta_i{}^2 x_h{}^{i2} - \zeta_i{}^1 x_h{}^{i1} + b_{hl}w^l = 0 && (h, l = 1, \cdots, r), \end{aligned}$$

where $L_i(\eta, \mu)$, Φ_β, ζ_i are the functions appearing in equations (4.4) and (4.5). By a *characteristic root* is meant a value of σ for which there exists a solution (η, μ, w) of these equations having continuous derivatives $\dot{\eta}^i$, $\ddot{\eta}^i$, $\dot{\mu}^\beta$ and with $(\eta) \neq (\tau\dot{x})$. In view of our remarks in Section 4 and the choice of the functions K_{ij} it is clear that the system (6.2) is independent of the coördinate system (x) used. It follows that the characteristic roots σ are invariants.

The following theorem can be established by the usual method (See, for example, VI, p. 524).

THEOREM 6.1. *If the second variation $J_2(\eta, w)$ of the functional J is non-negative along g for every set of admissible variations (η, w), then there can be no negative characteristic roots of the accessory boundary value problem* (6.2).

The end conditions (1.3) will be said to be *regular* on g if the matrix obtained by deleting the last two columns of the matrix

$$(6.3) \qquad \left\| \begin{matrix} x_h{}^{i1} & \dot{x}^i(t^1) & 0 \\ x_h{}^{i2} & 0 & \dot{x}^i(t^2) \end{matrix} \right\|$$

has rank r. The arc g will be said to satisfy the *non-tangency condition* in case the last two columns are not linearly dependent on the first r columns. The matrix (6.3) accordingly has rank $r + 2$ if and only if the end conditions are regular and the non-tangency condition holds on g.

The extremal g will be said to satisfy the *Clebsch S-condition* if the relation (3.4) with the equality sign excluded holds along g for every set $(\pi) \neq (\tau\dot{x})$ satisfying the equations (3.5).

We can now prove the theorem:

THEOREM 6.2. *Suppose the end conditions are regular and the non-tangency condition holds on the extremal g. If g is non-singular, then the second variation $J_2(\eta, w)$ of J is positive (non-negative) for every set of admissible variations $(\eta, w) \neq (\tau\dot{x}, 0)$ if and only if the Clebsch S-condition holds and all the characteristic roots of the accessory boundary value problem are positive (non-negative).*

To prove this we first note that under the transformation (5.10) the system of equation (6.2) are equivalent to the system

$$
\begin{aligned}
&M_p{}^0(u, \rho, \sigma) = L_p{}^0(u, \rho) + \sigma K_{pq}{}^0 u^q = 0, && \Phi_\beta{}^0 = 0,\\
(6.4) \qquad &u^{ps} - c_h{}^{ps} w^h = 0 && (p, q = 1, \cdots, n;\ s = 1, 2),\\
&v_i{}^2 c_h{}^{i2} - v_i{}^1 c_h{}^{i1} + b_{hl}{}^0 w^l = 0 && (h, l = 1, \cdots, r),
\end{aligned}
$$

where the functions $L_p{}^0(u, \rho)$, $\Phi_\beta{}^0$, v_p, $c_h{}^{i1}$, $c_h{}^{i2}$ are those appearing in equations (5.8), (5.9) and

$$K_{pq}{}^0 = b_p{}^i K_{ij} b_q{}^j, \qquad K_{ij} = a_i{}^p K_{pq}{}^0 a_j{}^q.$$

We may suppose further that the functions $b_p{}^i(x)$ have been chosen so that $K_{pq}{}^0$ is equal to the Kronecker delta along g. The system of equations (6.4) is then the boundary value problem usually associated with the functional $J_2{}^0(u, w)$. Its characteristic roots are defined to be the values of σ for which there exists a solution (u, ρ, w) having continuous derivatives $\dot{u}^p$, $\ddot{u}^p$, $\dot{\rho}^\beta$ and with $(u) \neq (0)$. It is clear that the functions M_i, $M_p{}^0$ satisfy equations (5.12) with L replaced by M. It follows that a value σ is a characteristic root of the boundary value problem (6.2) if and only if it is a characteristic root of the system (6.4). Moreover the analogue of Theorem 6.2 is true for the functional $J_2{}^0(u, w)$ in the class of normal admissible variations (u, w), as has been shown by the author (XIII, pp. 812-813). The theorem now follows from Lemma 5.1.

7. Sufficient conditions for relative minima. In this section we shall show that sufficient conditions as well as necessary conditions can be derived with the help of known theorems for the non-parametric problem. To this end we introduce the notion of a Mayer field.

A *Mayer field* is a region $\mathfrak{F}$ in $\mathfrak{R}$ together with a set of slope functions $r^i(x)$ and multipliers $\lambda^\beta(x)$ with the following properties:

(a) The slope functions $r^i(x)$ are the contravariant components of a vector of class C' and with $(r) \neq (0)$. The multipliers $\lambda^\beta(x)$ are invariant functions of class C'.

(b) The elements (x, r) satisfy the equations $\phi_\beta = 0$.

(c) The Hilbert integral

$$I^* = \int F_{r^i}(x, r, \lambda)\, dx^i$$

formed for these functions and $\lambda^0 = 1$ is independent of the path in $\mathcal{F}$.

It is well known that a solution $x^i = x^i(t)$ of the equations $\dot{x}^i - r^i(x) = 0$ together with the multipliers $\lambda^\beta[x(t)]$ define an extremal, called an *extremal of the field* $\mathcal{F}$.

If g is an extremal of a field $\mathcal{F}$ then the conditions (3.2) for g can be written in the form

$$[dI^*]_1{}^2 + d\theta \equiv 0. \tag{7.1}$$

In view of this fact it follows readily that the second differential

$$[d^2I^*]_1{}^2 + d^2\theta \tag{7.2}$$

is a quadratic form in $d\alpha_h$ whenever the relation (7.1) holds identically in $d\alpha_h$. The Weierstrass E-function $E(x, r, \lambda, \dot{x})$ is defined by equations (3.3).

By an argument like that used by Hestenes in the non-parametric case (XIII, pp. 805-806) we can prove the following theorem:

THEOREM 7.1. *If g is an extremal of a Mayer field at each point of which the inequality*

$$E[x, r(x), \lambda(x), \dot{x}] > 0 \tag{7.3}$$

holds for every set $(x, \dot{x}) \neq (x, kr)$ $(k > 0)$ satisfying equations $\phi_\beta = 0$ and if the endpoints of g and the set $(\alpha) = (0)$ are such that the equation (7.1) is an identity in $d\alpha_h$ and the quadratic form (7.2) is positive definite, then the arc g affords a proper minimum to the functional J relative to admissible arcs C in $\mathcal{F}$ with sets (α) near $(\alpha) = (0)$.

An extremal arc will be said to satisfy the *Weierstrass S-condition* if at each element $(x, \dot{x}, \lambda)$ on it the inequality

$$E(x, \dot{x}, \lambda, r) > 0 \tag{7.4}$$

holds for every set $(x, r) \neq (x, k\dot{x})$ $(k > 0)$ satisfying the equations $\phi_\beta = 0$. The *Clebsch S-condition* has been defined in the paragraph preceding Theorem 6.2.

LEMMA 7.1. *If an extremal g satisfies the Weierstrass and Clebsch S-conditions then the inequality (7.4) holds for all sets $(x, \dot{x}, \lambda)$ in a neigh-*

borhood of those belonging to g and for every set $(x, r) \neq (x, k\dot{x})$ $(k > 0)$ *satisfying the equations* $\phi_\beta = 0$.

The proof is well known and can be made by the arguments like those used by Morse (X, p. 121) for the case in which there are no differential side conditions.

We come now to the following important result:

THEOREM 7.2. *In order that an admissible arc g without corners and not intersecting itself afford a proper strong relative minimum to the functional J it is sufficient that there exists a set of multipliers $\lambda^0 = 1$, $\lambda^\beta(x)$ with which g satisfies the Euler-Lagrange equations* (3.1), *the transversality condition* (3.2), *the Weierstrass S-condition, the Clebsch S-condition, and that the second variation* (4.1) *of J be positive along g for every set of admissible variations* $(\eta, w) \neq (\tau\dot{x}, 0)$.

In order to prove this theorem we shall construct an auxiliary non-parametric problem of Bolza as follows: Let a neighborhood of g be represented by a single admissible coördinate system (x) with $\dot{x}^0(t) > 0$ along as described in Lemma 2.1. The coördinates $(x^0, x^1, \cdots, x^n)$ will also be termed $(x, y_1, \cdots, y_n)$. Since the Clebsch S-condition implies that g is a non-singular extremal, the function $\dot{x}^0(t)$ is of class C^3 for admissible parameters t. We can accordingly choose $t = x^0 = x$ as the parameter t defining g. The equations for g can then be written in the form

$$y_p = y_p(x) \qquad (x_1 \leqq x \leqq x_2;\ p = 1, \cdots, n). \tag{7.5}$$

Moreover the functions ϕ, f can be written in the form

$$\begin{aligned} \phi_\beta(x, y_1, \cdots, y_n, 1, y'_1, \cdots, y'_n) &= \phi_\beta{}^0(x, y, y'), \\ f(x, y, \cdots, y, 1, y', \cdots, y') &= f^0(x, y, y'). \end{aligned}$$

This suggests the problem of minimizing the functional

$$J^0 = \theta(\alpha) + \int_{x_1}^{x_2} f^0(x, y, y')\,dx$$

in the class of arcs (7.5) and sets (α) which satisfy the differential equations and end conditions

$$\begin{aligned} &\phi_\beta{}^0(x, y, y') = 0 && (\beta = 1, \cdots, m < n), \\ &x_s = x_s(\alpha), \quad y_p(x_s) = y_{ps}(\alpha) && (s = 1, 2), \end{aligned}$$

where

$$x_s(\alpha) = x^{0s}(\alpha), \quad y_{ps}(\alpha) = x^{ps}(\alpha) \qquad (p = 1, \cdots, n).$$

This problem is of the type already studied by the author. Clearly g is also

an extremal for this new problem, its multipliers being $\lambda^0 = 1$, $\lambda^\beta(x)$. The analogue of the Clebsch S-conditions also holds along g. Moreover if in Section 5 we select the functions $a_i{}^p$ so that for the coördinate system here used the equations (5. 1) take the form

$$u^p = \eta^p - \dot{x}^p(x^0)\eta^0 \qquad (p = 1, \cdots, n)$$

and the functions a_i so that $(a_0, a_1, \cdots, a_n) = (1, 0, \cdots, 0)$ then by setting $t = x$ it is found that the functional (5. 6) is the second variation of the functional J^0, that the equations (5. 3) are the equations of variation of the equations $\phi_\beta{}^0 = 0$, and that the equations (5. 4) are the secondary end conditions for the problem just described. It follows from Lemma 5. 1 that the second variation $J_2{}^0(u, w)$ is positive for every non-null set of admissible variations (u, w).

As a second step in the proof of Theorem 7. 2 we note that it is sufficient to prove the theorem for the case in which the end conditions (1. 3) and the function $\theta(\alpha)$ are of the form

$$x^{i1} = x^{i1}(\alpha^1, \cdots, \alpha^\rho), \qquad x^{i2} = x^{i2}(\alpha^{\rho+1}, \cdots, \alpha^r),$$
$$\theta(\alpha) = \theta^1(\alpha^1, \cdots, \alpha^\rho) - \theta^2(\alpha^{\rho+1}, \cdots, \alpha^r).$$

The proof of this statement is like that given by the author for the non-parametric case (XIII, pp. 815-816). The end conditions for the auxiliary non-parametric described in the preceding paragraph will also be in this separated form. It follows from the Clebsch S-condition and the positiveness of the second variation that there exists for this non-parametric problem a Mayer field $\mathfrak{F}$ such that the analogue of equation (7. 1) holds and that the quadratic form analogous to (7. 2) is positive definite. But when these conditions are interpreted in terms of the parametric problem it is found that $\mathfrak{F}$ is also a Mayer field for the parametric problem and that the condition (7. 1) holds and that the quadratic form (7. 2) is positive definite. By taking $\mathfrak{F}$ sufficiently small the condition (7. 3) will also be satisfied by virtue of the Weierstrass S-condition and Lemma 7. 1. Theorem 7. 2 now follows Theorem 7. 1.

Combining Theorems 6. 2 and 7. 2 we obtain the following result:

THEOREM 7. 3. *Suppose the end conditions are regular and the non-tangency condition holds on an admissible arc g without corners and not intersecting itself. If there exists a set of multipliers $\lambda^0 = 1$, $\lambda^\beta(t)$ with which g satisfies the Euler-Lagrange equations* (3. 1), *the transversality condition* (3. 2), *the Weierstrass S-condition, the Clebsch S-condition, and if all the*

characteristic roots of the accessory boundary value problem are positive, then g affords a proper strong relative minimum to J.

Similarly the analogues of the sufficient conditions of "Theorems 9.1, 9.3, 9.4" of Hestenes (XIII, cf. XVI) and also those of Morse (XIV) and Reid (XV) can be established without difficulty.

HARVARD UNIVERSITY.

REFERENCES.

I. Bolza, *Vorlesungen über Variationsrechnung*, 1909.

II. Bliss, "Jacobi's condition for problems of the calculus of variations in parametric form," *Transactions of the American Mathematical Society*, vol. 17 (1916), pp. 195-206.

III. Eshleman, *The Lagrange Problem in Parametric Form*, Dissertation, The University of Chicago, 1922.

IV. Graves, "Discontinuous solutions in space problems of the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 1-28.

V. Bliss, "The problem of Lagrange in the calculus of variations," *American Journal of Mathematics*, vol. 52 (1930), pp. 673-744.

VI. Morse, "Sufficient conditions in the problem of Lagrange with variable end points," *American Journal of Mathematics*, vol. 53 (1931), pp. 517-596.

VII. Hefner, "The condition of Mayer for discontinuous solutions of the Lagrange problem," *Contributions to the Calculus of Variations* (1931-2), The University of Chicago Press, pp. 95-130.

VIII. Graves, "On the Weierstrass condition for the problem of Bolza in calculus of variations," *Annals of Mathematics* (2), vol. 33 (1932), pp. 747-752.

IX. Tucker, "On tensor invariance in the calculus of variations," *Annals of Mathematics* (2), vol. 35 (1934), pp. 341-350.

X. Morse, "The calculus of variations in the large," *American Mathematical Society Colloquium Publications*, vol. 18, New York (1934).

XI. Reid, "Analogues of the Jacobi condition for the problem of Mayer in the calculus of variations," *Annals of Mathematics*, vol. 35 (1934), pp.

XII. Hestenes, "A note on the Jacobi condition for parametric problems in the calculus of variations," *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 297-302.

XIII. Hestenes, "Sufficient conditions for the problem of Bolza in the calculus of variations," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 793-818.

XIV. Morse, "Sufficient conditions in the problem of Lagrange without assumptions of normalcy," *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 147-160.

XV. Reid, "The theory of the second variation for the non-parametric problem of Bolza," *American Journal of Mathematics*, vol. 57 (1935), pp. 573-586.

XVI. Hestenes, "On sufficient conditions in the problems of Lagrange and Bolza." To appear soon in the *Annals of Mathematics*.

ON EXPANSIONS IN TERMS OF A CERTAIN GENERAL CLASS OF FUNCTIONS.

By W. T. MARTIN.

Introduction. The object of this paper is to define certain generalizations of Appell polynomials and to study the problem of expanding a given function in a series of these generalized functions. A set of *Appell polynomials* $\{P_n(x)\}$ is defined by $e^{xz}\sum_0^\infty p_n z^n \sim \sum_0^\infty P_n(x)z^n$ where the series $\Sigma p_n z^n$ may or may not have a region of convergence. The functions $\{P_n(x)\}$ are such that $P_n(x)$ is a polynomial of degree n in x and $P'_n(x) = P_{n-1}(x)$, $n > 0$. Polynomials of this character were introduced by Appell and they, together with certain generalizations of them, have been studied by various writers.[1] When $\Sigma p_n z^n = z/(e^z - 1)$, the polynomials $\{n!\,P_n(x)\}$ are the *Bernoulli polynomials* $\{B_n(x)\}$. The *Bernoulli-Hurwitz functions* $B_{n,r}(x)$ $(n, r = 0, 1, 2, \cdots)$, are defined by

$$ze^{xz}/(e^z - 1) = \sum_{n=-\infty}^{\infty} B_{n,r}(x)z^n/n!, \qquad 2\pi r < |z| < 2\pi(r+1).$$

Carmichael has developed the theory of the expansion of functions in series of Bernoulli polynomials and of Bernoulli-Hurwitz functions.[2] Many of the ideas of the present paper were suggested by a study of his paper.

Replacing the function e^{xz} by a function $E(x,z)$ of a given character and denoting by $g(z)$ an analytic function having specified zeros, we define functions $A_{n,r}(x)$ by means of the Laurent's developments of $E(x,z)/g(z)$ in annular rings. In part I we study convergence properties of series of the sort $\sum_0^\infty c_n A_{n,r}(x)$, where r is a fixed non-negative integer.

In part II we study the properties of the functions represented by convergent series of the above sort and determine necessary and sufficient conditions on a given function in order that it be expressible in a series of this sort.

---

[1] P. Appell, *Ann. Sci. Norm. Sup.*, ser. 2, vol. 9 (1880), pp. 119-144; H. Léauté, *Journal de Mathématiques*, ser. 3, vol. 7 (1881), pp. 185-200; G. H. Halphen, *Bulletin des Sciences Mathématiques*, ser. 2, vol. 5 (1881), pp. 462-468; A. Angelesco, *Comptes Rendus*, vol. 176 (1923), pp. 275-278; I. M. Sheffer, *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 261-280; I. M. Sheffer, *American Journal of Mathematics*, vol. 53 (1931), pp. 15-38.

[2] R. D. Carmichael, *Annals of Mathematics*, ser. 2, vol. 34 (1933), pp. 349-378.

In part III we introduce an operator B which is such that $BA_{n,r}(x) = A_{n-1,r}(x)$. The problem of expanding a given function in a series of the above sort may also be viewed as a problem of solving a certain functional equation. We give existence theorems for these associated functional equations. When $E(x,z) = e^{xz}$ the function $BF(x)$ is the derivative $F'(x)$ of functions $F(x)$ of finite exponential type. For this case the theory furnishes solutions of exponential type of differential equations of infinite order, yielding results previously obtained by Pincherle and Perron.[3] As a corollary we obtain the solutions of a class of infinite systems of linear equations in an infinite number of unknowns treated previously (by different methods) by Perron.[4] When $E(x,z) = (1+z)^x$ the function $BF(x)$ is the difference $\Delta F(x) \equiv F(x+1) - F(x)$ of a certain class of integral functions.

I. Convergence Theory.

1. 1. Definition of the functions $A_{n,r}(x)$. Let $u_0(x), u_1(x), \cdots$, be any infinite sequence of integral functions possessing the following two properties:

1°. $\limsup_{n=\infty} | u_n(x) |^{1/n} \leqq 1/f \ (0 < f \leqq \infty)$, for all x;

2°. if a function $F(x)$, representable by a series of the form $\sum_0^\infty e_n u_n(x)$ for which

$$\limsup_{n=\infty} | e_n |^{1/n} < f,$$

is identically zero, then $e_n = 0$, $n = 0, 1, \cdots$.

Examples of such sequences with interesting connections are:

$$\{x^n/n!\}, \qquad \{x^n/(\lambda_0\lambda_1 \cdots \lambda_n)\}, \qquad \{x(x-1)\cdots(x-n+1)/n!\},$$
$$\{x(x-\lambda_1)\cdots(x-\lambda_{n-1})/(\lambda_1\cdots\lambda_n)\},$$

where $\lambda_n \neq 0$, $\lim | \lambda_n | = \infty$ and in the fourth example $| \lambda_1 | \leqq | \lambda_2 | \leqq \cdots$.

Let us form the function $E(x,z) = \sum_0^\infty u_n(x) z^n$ which, we see, is analytic for all $| z | < f$ and for all x. Let $g(z) = \sum_0^\infty \beta_n z^n$, $\beta_0 \neq 0$, be an analytic function regular for $| z | < \rho$ $(0 < \rho \leqq f)$. Let $g(z)$ have zeros[5] of multiplicities s_m at points α_m, $| \alpha_m | < \rho$, $(m = 1, 2, \cdots)$.

---

[3] S. Pincherle, *Acta Mathematica*, vol. 48 (1926), pp. 279-304 (first published in 1888); O. Perron, *Mathematische Annalen*, vol. 84 (1921), pp. 31-42.

[4] O. Perron, *ibid.*, pp. 1-15. See in particular Satz 1.

[5] We will carry through the theory based upon the assumption that $g(z)$ has an

As a matter of notation we assume that the zeros α_m of $g(z)$ are arranged according to increasing absolute values and that those of the same absolute value are arranged according to increasing arguments, where the argument θ_m of α_m is taken to be in the range $0 \leqq \theta_m < 2\pi$. Let there be $k(1)$ zeros of smallest absolute value γ_1, $k(2)$ of next smallest absolute value γ_2, etc. Let us denote by $\mu(r)$ the sum $k(1) + \cdots + k(r)$. We define $\gamma_0 = k(0) = \mu(0) = 0$. Let C_r be a circle of radius C'_r, $\gamma_r < C'_r < \gamma_{r+1}$, about the origin as center.

Let us consider the expansions [6]

$$(1.1) \qquad \frac{E(x,z)}{g(z)} = \sum_{n=-\infty}^{\infty} A_{n,r}(x) z^n, \quad \gamma_r < |z| < \gamma_{r+1}, \quad (r = 0, 1, \cdots).$$

We see that the $A_{n,r}(x)$ are expressible as

$$(1.2) \qquad A_{n,r}(x) = \frac{1}{2\pi i} \int_{C_r} \frac{E(x,z)}{g(z)} \frac{dz}{z^{n+1}}.$$

Denoting by $R_n(x, \alpha_m)\alpha_m^{-n-1}$ the residue of $E(x,z)^{-n-1}/g(z)$ at $z = \alpha_m$, we see that $R_n(x, \alpha_m)$ has the form

$$(1.3) \qquad R_n(x, \alpha_m) = \sum_{k=0}^{s_m-1} \frac{(-1)^k}{\alpha_m^k} \binom{n+k}{k} p_{m,k}(D_m) E(x,z), \quad (n = 0, 1, \cdots),$$

where $p_{m,k}(t)$ is a polynomial of degree $s_m - 1 - k$ in t and D_m is an operator defined by

$$(1.4) \qquad D_m{}^j H(x,z) \equiv \frac{\partial^j}{\partial z^j} H(x,z) \Big|_{z=\alpha_m}.$$

1.2. Convergence theorems. Let us contemplate series of the form

$$(1.5) \qquad \sum_{n=0}^{\infty} c_n A_{n,r}(x),$$

where r is a fixed non-negative integer. Let us assume that this series converges [converges absolutely] for h points $x_1, \cdots, x_h$ for which the determinant Δ is non-vanishing. The determinant Δ contains h rows whose j-th one we indicate:

---

infinite number of zeros. No essential modification is necessary for the case where $g(z)$ has a finite number of zeros $\alpha_1, \cdots, \alpha_R$, $R \geqq 1$.

[6] The functions $P_{n,r}(x)$, $(n, r = 0, 1, \cdots)$, generated by $e^{xz}/g(z)$ are Appell polynomials $(r = 0)$ and associated Appell functions. Bird has introduced the functions $P_{-n,r}(x)$, $(n, r = 1, 2, \cdots)$, as the coefficients of the Laurent's expansion of $e^{xz}/g(z)$ in a certain annular ring. See M. T. Bird, *On generalizations of sum formulas of the Euler-Maclaurin type*, Illinois dissertation (1934).

$$\Delta \equiv | \delta^0_{\mu(r)+1,j} \cdots \delta^{\xi(1)}_{\mu(r)+1,j} \; \delta^0_{\mu(r)+2,j} \cdots \delta^{\xi(2)}_{\mu(r)+2,j} \cdots \delta^0_{\mu(r+1),j} \cdots \delta^{\xi(k(r+1))}_{\mu(r+1),j} |,$$

where

$$h = \sum_{m=\mu(r)+1}^{\mu(r+1)} s_m, \qquad \delta^k_{m,j} \equiv D_m{}^k E(x_j, z), \qquad \xi(q) + 1 = s_{\mu(r)+q}.$$

Let us assume temporarily that

$$\limsup_{n=\infty} | c_n |^{1/n} \leqq \gamma_{r+1}. \tag{1.6}$$

From (1.1) we see that

$$\limsup_{n=\infty} | A_{n,r}(x) |^{1/n} = 1/\gamma_{r+1}, \tag{1.7}$$

and hence in view of (1.6) each of the series

$$\sum_{n=0}^{\infty} c_n A_{n,r+p}(x), \qquad (p = 1, 2, \cdots), \tag{1.8}$$

converges absolutely and uniformly in every finite region. Forming (1.8) for $p = 1$ and subtracting from it the series (1.5), we see that the difference

$$\sum_{n=0}^{\infty} \frac{c_n}{\gamma^{n+1}_{r+1}} \sum_{m=\mu(r)+1}^{\mu(r+1)} e^{-i(n+1)\theta_m} \sum_{k=0}^{s_m-1} \frac{(-1)^k}{\alpha_m{}^k} \binom{n+k}{k} p_{m,k}(D_m) E(x, z) \tag{1.9}$$

converges [converges absolutely] for $x = x_1, \cdots, x_h$.

The determinant Δ_1 obtained from Δ by replacing $\delta^k_{m,j}$ by $p_{m,k}(D_m) E(x_j, z)$ is non-vanishing since it is obtainable from Δ by means of a finite number of elementary transformations of determinants. Since Δ_1 is different from zero the convergence [absolute convergence] of the series (1.9) for $x = x_1, \cdots, x_h$ implies the convergence [absolute convergence] of each of the series

$$\sum_{n=0}^{\infty} \frac{c_n}{\gamma^{n+1}_{r+1}} \binom{n+k}{k} e^{-i(n+1)\theta_m} \tag{1.10}$$
$$(k = 0, \cdots, s_m - 1;\ m = \mu(r) + 1, \cdots, \mu(r+1)).$$

Hence the series (1.9) converges [converges absolutely] and uniformly in every finite region. Using the absolute and uniform convergence of the series (1.8) for $p = 1$, we obtain uniform [absolute and uniform] convergence of the series (1.5) in every finite region.

The preceding work depends upon the assumption that (1.6) holds. We show now that (1.6) must hold if the series (1.5) converges at the h points

$x_1, \cdots, x_h$. Assume for the moment that (1.6) does not hold.[7] Let η be a constant such that

$$\gamma_{r+1} < \eta < \gamma_{r+2}, \qquad \eta < \limsup_{n=\infty} | c_n |^{1/n}.$$

Consider the series (1.5) formed with c_n replaced by $c_n \zeta_n$ where

$$\zeta_0 = 1, \quad \zeta_n = 1/n^2 \text{ if } | c_n | \leqq \eta^n, \quad \zeta_n = \eta^n/(n^2 | c_n |) \text{ if } | c_n | > \eta^n.$$

It surely converges absolutely for $x = x_1, \cdots, x_h$ and furthermore we have

$$\limsup_{n=\infty} | c_n \zeta_n |^{1/n} = \eta.$$

From these facts we see that the series (1.8) formed with c_n replaced by $c_n \zeta_n$ converge absolutely and uniformly in every finite region. Reasoning as before, we have convergence for each of the series (1.10) with c_n replaced by $c_n \zeta_n$. But

$$\lim_{n=\infty} | c_n \zeta_n \gamma_{r+1}^{-n-1} | \neq 0.$$

So we are led to a contradiction and (1.6) must hold.

We have proved the following theorem and corollary:

THEOREM 1.1. *If the series* (1.5) *converges* [*converges absolutely*] *for* h *points* $x_1, \cdots, x_h$ *for which the determinant* Δ *is non-vanishing, then in every finite region of the x-plane each of the series* (1.5) *and* (1.8) *converges* [*converges absolutely*] *and uniformly.*

COROLLARY 1.1. *Under the hypotheses of the preceding theorem the series*

$$\phi(z) = \sum_{n=0}^{\infty} c_n z^{-n-1} \tag{1.11}$$

converges [*converges absolutely*] *for* $z = \alpha_m$ $(m = \mu(r) + 1, \cdots, \mu(r+1))$, *and hence the series* (1.11) *converges absolutely for* $| z | > \gamma_{r+1}$. *Indeed, each of the series* (1.10) *converges* [*converges absolutely*].

Let R be any finite region of the x-plane. We want to show that there exists in R a set of h points $x_1, \cdots, x_h$ for which Δ is non-vanishing. We use a special case of the following lemma:

LEMMA 1.1. *If* $a_1, \cdots, a_p$ *are any distinct non-zero constants,* $| a_k | < f$ $(k = 1, \cdots, p)$, *and if*

$$\sum_{k=1}^{p} \sum_{s=0}^{q} c_{sk} (\partial^s / \partial z^s) E(x, z) |_{z=a_k} \equiv 0,$$

then $c_{sk} = 0$, $(s = 0, \cdots, q;\ k = 1, \cdots, p)$.

[7] The fundamental ideas of this argument are due to H. J. Miles. I am also indebted to Miles for much of the notation of this part.

The truth of this lemma is an immediate consequence of the property 2° of the functions $\{u_n(x)\}$. Using this lemma and making an easy induction argument, we obtain the following lemma:

LEMMA 1. 2. *In every region R of the x-plane there exists a set h points* $x_1, \cdots, x_h$ *for which the determinant* Δ *is non-vanishing.*

Theorem 1. 2 follows from the preceding lemma.

THEOREM 1. 2. *The hypotheses of Theorem* 1. 1 *are satisfied whenever there is any region in which the series* (1. 5) *converges* [*converges absolutely*].

II. Expansions in Terms of the Functions $A_{n,r}(x)$.

2. 1. A class of integral functions. Let $F(x)$ be any integral function enjoying the property that it has an expansion of the form

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x) \tag{2.1}$$

with

$$\limsup_{n=\infty} |e_n|^{1/n} = q, \qquad 0 \leqq q < f. \tag{2.2}$$

We shall, in such a case, say that $F(x)$ is of *sort* $\{u_n\}$, *type* q. The proof of the following lemma relating to functions of sort $\{u_n\}$, type q is immediate.

LEMMA 2. 1. *A necessary and sufficient condition that a function* $F(x)$ *be of sort* $\{u_n\}$, *type not exceeding* q, *is that it be expressible in the form*

$$F(x) = (1/2\pi i) \int_C E(x,z)\, a(z)\, dz,$$

where $a(z)$ *is analytic for* z *in the annular ring* $q < |z| \leqq s \leqq \infty$ *and* C *is a circle of radius* $q + \epsilon < s$ *about the origin as center and* ϵ *is a positive constant such that* $q + \epsilon < f$.

2. 2. Properties of series of functions $A_{n,r}(x)$. Suppose we have a series

$$F(x) = \sum_{n=0}^{\infty} c_n A_{n,r}(x), \tag{2.3}$$

where r is a fixed non-negative integer, satisfying the conditions of Theorem 1. 1 as to convergence. We seek to determine the class of functions $F(x)$ represented by such convergent series. Let us set

$$\limsup_{n=\infty} |c_n|^{1/n} = q. \tag{2.4}$$

Then, as we have seen, $q \leqq \gamma_{r+1}$ and the function $\phi(z)$ defined by (1.11) is analytic for $|z| > q$. Let us form the function

$$H(x) = \frac{1}{2\pi i} \int_{C_{r+1}} \frac{E(x, z)}{g(z)} \phi(z)\, dz \tag{2.5}$$

which by Lemma 2.1 is of sort $\{u_n\}$, type not exceeding γ_{r+1}. Integrating term by term the expression obtained by substituting the expansion (1.11) of $\phi(z)$ into (2.5), we see that

$$H(x) = \sum_{n=0}^{\infty} c_n A_{n,r+1}(x). \tag{2.6}$$

The difference $H(x) - F(x)$ is given by the expression (1.9). By Corollary 1.1 the convergence of the series (2.3) implies the convergence of each of the series (1.10) and hence we may separate the sum in (1.9) into a finite number of terms of the sort

$$D_m{}^k E(x, z) \sum_{n=0}^{\infty} \frac{c_n}{\gamma_{r+1}^{n+1}} \binom{n+k}{k} e^{-i(n+1)\theta_m}.$$

From the form of $D_m{}^k E(x, z)$ we see that it is of sort $\{u_n\}$, type $|\alpha_m|$. Thus the function $F(x)$, being expressible as a finite sum of functions each of sort $\{u_n\}$, type not exceeding γ_{r+1}, is itself of sort $\{u_n\}$, type not exceeding γ_{r+1}.

THEOREM 2.1. *If a function $F(x)$ has the convergent expansion (2.3) then $F(x)$ is of sort $\{u_n\}$, type not exceeding γ_{r+1}.*

Expressing the series in (2.3) as a series of the form (2.1) and using property 2° of the sequence $\{u_n(x)\}$, we have the following theorem:

THEOREM 2.2. *No function $F(x)$ can have two distinct expansions in series of the form (2.3) for the same value of r.*

Let $F(x)$ be a given function of sort $\{u_n\}$, type q, where $\gamma_r \leqq q < \gamma_{r+1}$. We study the problem of expanding $F(x)$ in a series of the form (2.3). The function $F(x)$ does have an expansion of the form (2.1) with (2.2) holding and hence we may write

$$F(x) = \frac{1}{2\pi i} \int_{C'_r} \frac{E(x, z)}{g(z)}\, g(z) \sum_{n=0}^{\infty} e_n z^{-n-1} dz, \tag{2.7}$$

where $C'_r = q + \epsilon < \gamma_{r+1}$ and ϵ is positive. Writing

$$g(z) \sum_{n=0}^{\infty} e_n z^{-n-1} = \sum_{n=-\infty}^{\infty} c_n z^{-n-1} = \sum_{n=0}^{\infty} c_n z^{-n-1} + P(z), \qquad q < |z| < \rho,$$

we have

$$(2.8)\qquad F(x)=\frac{1}{2\pi i}\int_{C_r}\frac{E(x,z)}{g(z)}\,[P(z)+\sum_{n=0}^{\infty}c_n z^{-n-1}]dz.$$

Integrating term by term, we have (see (1.2))

$$(2.9)\qquad F(x)=\sum_{m=1}^{\mu(r)}\sum_{k=0}^{s_m-1}d_{mk}D_m{}^kE(x,z)+\sum_{n=0}^{\infty}c_nA_{n,r}(x),$$

where the c's and d's are uniquely determined constants. If $r=0$, the finite sum on the right of (2.8) vanishes since C_0 contains no zero of $g(z)$.

THEOREM 2.3. *Any given function $F(x)$ of sort $\{u_n\}$, type q, such that $\gamma_r\leqq q<\gamma_{r+1}$ has an expansion of the form (2.9) where the c's and d's are uniquely determined constants.*

We may also write $F(x)$ in the form (2.8) with C_r replaced by C_{r+p}, where p is any positive integer. Since the Laurent's series in (2.8) is valid for z in an annular ring which includes the circle C_{r+p} we have the following corollary:

COROLLARY 2.1. *The function $F(x)$ has an expansion in the form (2.9) with r replaced throughout by $r+p$, where p is any positive integer. The constants c_n $(n=0,1,\cdots)$, and d_{mk} $(k=0,\cdots,s_m-1;\ m=1,\cdots,\mu(r))$, are the same in every case.*

Suppose $F(x)$ is a given function of sort $\{u_n\}$, type γ_{r+1}. Then, by Theorem 2.3, it has an expansion of the form

$$(2.10)\qquad F(x)=\sum_{m=1}^{\mu(r+1)}\sum_{k=0}^{s_m-1}d_{mk}D_m{}^kE(x,z)+\sum_{n=0}^{\infty}c_nA_{n,r+1}(x).$$

If $F(x)$ also has an expansion of the form (2.9) it is easily seen that the coefficients c_n be the same in the two cases. Since the series (1.9) represents the difference $\Sigma c_nA_{n,r+1}(x)-\Sigma c_nA_{n,r}(x)$ the expansions (2.9) and (2.10) can coexist only if the series (1.9) converges. In section 1.2 we proved that the convergence of the series (1.9) implies the convergence of each of the series (1.10). From its form we see that if (1.9) converges it gives the negative of the terms for $m=\mu(r)+1,\cdots,\mu(r+1)$ in the finite sum in (2.10); furthermore, when the series (1.9) converges the two expansions (2.9) and (2.10) coexist.

This completes the proof of the following theorem:

THEOREM 2.4. *Let $F(x)=\Sigma e_nu_n(x)$ be any given function of sort $\{u_n\}$, type γ_{r+1}. Write*

$$\frac{1}{2\pi i}\int_{C_{r+1}} E(x,z)\,g(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz = \sum_{n=0}^{\infty} c_n u_n(x).$$

Then a necessary and sufficient condition that $F(x)$ have an expansion of the form (2.9) is that each of the series (1.10) shall converge.

These theorems give necessary and sufficient conditions for the expansion of functions in series of the functions $A_{n,r}(x)$.

III. Functional Equations.

3.1. The operator B. We define the operator B to be such that if $F(x)$ be any given function of sort u_n, type less than f, that is, if

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x), \qquad \limsup_{n=\infty} |e_n|^{1/n} < f, \tag{3.1}$$

then

$$B^m F(x) = \sum_{n=0}^{\infty} e_{n+m} u_n(x) \qquad (m = 0, 1, \cdots). \tag{3.2}$$

Expanding $A_{n,r}(x)$ $(r = 0, 1, \cdots;\ n = 0, \pm 1, \pm 2, \cdots)$, in series of the form (3.1), we see that $B\,A_{n,r}(x) = A_{n-1,r}(x)$.

3.2. Existence theorems. Let $h(z) = \sum_{0}^{\infty} b_n z^n$, $b_0 \neq 0$, be a series convergent for $|z| < \sigma$ where $0 < \sigma \leqq f$. Let q be any positive number less than σ. Let $\theta(x) = \sum_{0}^{\infty} c_n u_n(x)$ be any given function of sort $\{u_n\}$, type not exceeding q. Let us seek solutions

$$F(x) = \sum_{n=0}^{\infty} e_n u_n(x) \tag{3.3}$$

of the linear functional equation

$$b_0 F(x) + b_1 BF(x) + b_2 B^2 F(x) + \cdots = \theta(x), \tag{3.4}$$

under the following hypothesis on the coefficients:

$$\limsup_{n=\infty} |e_n|^{1/n} \leqq q. \tag{3.5}$$

The trial solution (3.3) may be written in the form

$$F(x) = \frac{1}{2\pi i}\int_C E(x,z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz, \tag{3.6}$$

where C is any circle about the origin as center of radius $q + \epsilon < \sigma$ and ϵ is a positive number such that $h(z)$ has no zero in the ring $q < |z| \leqq q + \epsilon$.

Substituting the trial solution (3. 6) into the equation (3. 4) and performing an interchange of operations, we see that a necessary and sufficient condition that (3. 3) or its equivalent (3. 6) be a solution of (3. 4) is that

$$(3.7) \qquad \theta(x) = \frac{1}{2\pi i}\int_C E(x,z)h(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz,$$

which holds if and only if the e's satisfy the conditions

$$(3.8) \qquad \sum_{\mu=0}^{\infty} b_\mu e_{\mu+n} = c_n \qquad (n = 0, 1, \cdots).$$

A particular solution of (3. 4) of sort $\{u_n\}$, type not exceeding q is obviously given by

$$(3.9) \qquad F_1(x) = \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\sum_{n=0}^{\infty} c_n z^{-n-1}dz,$$

and hence a set of solutions $\{e_n\}$ of (3. 8) of the character described by (3. 5) is given by setting $e_n = d_n$ $(n = 0, 1, \cdots)$, where the d_n are defined by

$$[h(z)]^{-1}\sum_{n=0}^{\infty} c_n z^{-n-1} = \sum_{n=0}^{\infty} d_n z^{-n-1} + \sum_{n=0}^{\infty} f_n z^n, \qquad q < |z| \leqq q + \epsilon.$$

If we write the trial solution (3. 3) (or (3. 6)) in the form

$$\begin{aligned} F(x) &= \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\, h(z)\sum_{n=0}^{\infty} e_n z^{-n-1}dz \\ &= \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\, [\sum_{n=0}^{\infty} \delta_n z^{-n-1} + \sum_{n=0}^{\infty} \epsilon_n z^n]dz, \end{aligned}$$

then we see that $F(x)$ is a solution of (3. 4) if and only if $\delta_n = c_n$. In such a case

$$F(x) = F_1(x) + \frac{1}{2\pi i}\int_C \frac{E(x,z)}{h(z)}\sum_{n=0}^{\infty} \epsilon_n z^n dz.$$

If $h(z)$ has no zero for $|z| \leqq q$, then $F_1(x)$ is the only solution of (3. 4) of sort $\{u_n\}$, type not exceeding q. If $h(z)$ has zeros of multiplicities σ_m at points

$$(3.10) \qquad a_m, \qquad |a_m| \leqq q \qquad (m = 1, \cdots, k),$$

then by the theory of residues it follows that $F(x)$ must be of the form

$$(3.11) \qquad F(x) = F_1(x) + \sum_{m=1}^{k}\sum_{j=0}^{\sigma_m - 1} \frac{\partial^j}{\partial z^j} E(x,z)\big|_{z=a_m}.$$

Every solution of (3. 4) of sort $\{u_n\}$, type not exceeding q, is expressible in the form (3. 11) where the d's are constants. But each of the functions

$$(3.12) \qquad \frac{\partial^j}{\partial z^j} E(x,z)\,|_{z=a_m} = \sum_{n=j}^{\infty} n(n-1)\cdots(n-j+1)a_m^{n-j}u_n(x), \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k),$$

constitutes a solution of the homogeneous equation, that is, of equation (3.4) formed for $\theta(x) \equiv 0$. These solutions are linearly independent (see Lemma 1.1) and each is of sort $\{u_n\}$, type not exceeding q.

This completes the proof of the following theorem:

THEOREM 3.1. *Let* $h(z) = \sum_0^\infty b_n z^n$, $b_0 \neq 0$, *be a series convergent for* $|z| < \sigma \leqq f$. *Let* q *be any positive number less than* σ. *Let* $\theta(x) = \sum_0^\infty c_n u_n(x)$ *be any function of sort* $\{u_n\}$, *type not exceeding* q. *If* $h(z)$ *has no zero for* $|z| \leqq q$, *then the only solution of the equation* (3.4) *of sort* $\{u_n\}$, *type not exceeding* q, *is given by* (3.9). *If* $h(z)$ *has zeros of multiplicities* σ_m *at the points* (3.10) *then the most general solution of* (3.4) *of sort* $\{u_n\}$, *type not exceeding* q, *is given by* (3.11) *where the* d*'s are arbitrary constants.*[8]

For the case where $h(z) = g(z)$ the solutions $F(x)$ are expressible in the form (2.9) where r is the integer defined by $\gamma_r \leqq q < \gamma_{r+1}$.

Noting that every solution $F(x)$ of (3.4) of the desired character, when expressed in the form (3.3), furnishes a set of solutions $\{e_n\}$ of (3.8), verifying (3.5), we obtain as corollaries certain results relating to infinite systems of linear equations in an infinite number of unknowns. We state only the following one relating to homogeneous systems:[9]

COROLLARY 3.1. *Let* $h(z) = \sum_0^\infty b_n z^n$, $b_0 \neq 0$, *be a series convergent for* $|z| < \sigma$. *Let* q *be any positive number less than* σ. *If* $h(z)$ *has no zero for* $|z| \leqq q$, *then the only set of solutions* $\{e_n\}$ *of*

$$(3.13) \qquad \sum_{\mu=0}^{\infty} b_\mu e_{\mu+n} = 0, \qquad (n = 0, 1, \cdots),$$

verifying (3.5), *is the identically zero one. If* $h(z)$ *has* p *zeros (multiple zeros counted multiply),* $p > 0$, *for* $|z| \leqq q$, *then there exist exactly* p *linearly independent sets of solutions* $\{e_n\}$ of (3.13) *of the character described by*

---

[8] Since $\theta(x)$ may be identically zero the theorem applies also to homogeneous equations.

[9] Perron (*loc. cit.*) by different methods has obtained the results contained in this corollary.

(3.5). *If these p zeros are given by* (3.10), *where a_m is of multiplicity σ_m, then p such sets of solutions are given by*

(3.14) $$e_n = n(n-1)\cdots(n-j+1)a_m{}^{n-j} \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k),$$

or by a linear combination of these, namely

(3.15) $$e_n = n^j a_n{}^m \qquad (j=0,\cdots,\sigma_m-1;\ m=1,\cdots,k).$$

3.3. A particular case of the operator B. Let $\lambda_0, \lambda_1, \lambda_2, \cdots$, be an infinite sequence of non-zero constants such that

(3.16) $$\lim_{n=\infty} |\lambda_0\lambda_1\cdots\lambda_n|^{1/n} = \infty, \qquad \lim_{n=\infty} |\lambda_n|^{1/n} = 1.$$

The sequence $\{u_n(x) = x^n/(\lambda_0\lambda_1\cdots\lambda_n)\}$ forms a sequence possessing the two properties 1° and 2° in the introduction. We give here a definition of the operator B associated with this sequence which is an extension of the one already given. If $F(x)$ be any single-valued analytic function regular at $x=0$ then we define

(3.17) $$B^m F(x) = \frac{1}{2\pi i}\int_C \frac{F(z)}{z^{m+1}} \sum_{n=0}^{\infty} \frac{\lambda_0\lambda_1\cdots\lambda_{n+m}}{\lambda_0\lambda_1\cdots\lambda_n}\frac{x^n}{z^n}\,dz \quad (m=0,1,\cdots),$$

where C is any circle about the origin as center within the region of analyticity of $F(z)$ and x is any interior point of C. In view of (3.16) each of these integrals is defined for any point x interior to C and the Maclaurin's expansion of $B^mF(x)$ has the same circle of convergence as has the one representing $F(x)$. We note the following additional properties of the operator B:

$$Bx^n = \lambda_n x^{n-1}, \quad n>0; \qquad Bc = 0,\ c \text{ any constant.}$$

If

$$F(x) = \sum_{n=0}^{\infty} \delta_n \frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n}, \qquad |x| < R,$$

then

(3.18) $$B^mF(x) = \sum_{n=0}^{\infty} \delta_{n+m}\frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n}, \qquad |x| < R.$$ [10]

[10] Pincherle and Hadamard have studied operators of this sort. See S. Pincherle, *Giornale di Matematiche*, vol. 32 (1884), pp. 62-74; J. Hadamard, *Acta Mathematica*, vol. 22 (1899), pp. 55-63. Hadamard introduced and used a contour integral representation of the form (3.17) in his study of the *Hadamard product*. Carmichael suggested to me the above contour integral representation for the operation.

We may and will also consider (3.17) and (3.18) defined for m a negative integer provided that we agree to replace by zero any term which contains a λ or a δ with a negative subscript.

For the case $\lambda_0 = 1$, $\lambda_n = n$, $n > 0$, the operation $B\,F(x)$ furnishes the derivative of an analytic function regular at $x = 0$. For negative values of m, say $m = -k$, equations (3.17) and (3.18) represent generalizations of the multiple integrals

$$\int_0^x \int_0^{z_1} \cdots \int_0^{z_{k-1}} F(z_k)\, dz_k \cdots dz_2 dz_1.$$

We place a further restriction on the λ's, namely that

$$|\lambda_0| \leqq |\lambda_1| \leqq |\lambda_2| \leqq \cdots. \tag{3.19}$$

Carmichael[11] has obtained a general expansion theorem relating to a set of non-zero constants verifying (3.19) and such that $\lim |\lambda_n| = \infty$. I state it here for future reference.

General Expansion Theorem. *Let $a_1(x), a_2(x), \cdots$, be a finite or infinite sequence of functions each of which is analytic in the interior of the circle $|x| = \bar{\rho}$. Let a be a positive number less than $\bar{\rho}$. Denote by M_k the maximum absolute value of $a_k(x)$ in the region $|x| \leqq a$. Suppose the following series converges*

$$M_1 \frac{a}{|\lambda_1|} + M_2 \frac{a^2}{|\lambda_1\lambda_2|} + M_3 \frac{a^3}{|\lambda_1\lambda_2\lambda_3|} + \cdots. \tag{3.20}$$

Let $F(x)$ be analytic for $|x| \leqq a$. Then $F(x)$ has a unique expansion in either of the following two forms:

$$F(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n} \left\{ 1 + a_1(x) \frac{x}{\lambda_{n+1}} + a_2(x) \frac{x^2}{\lambda_{n+1}\lambda_{n+2}} + a_3(x) \frac{x^3}{\lambda_{n+1}\cdots\lambda_{n+3}} + \cdots \right\}, \tag{3.21}$$

$$F(x) = \sum_{n=0}^{\infty} c_n \frac{x^n}{\lambda_0\lambda_1\cdots\lambda_n} + a_1(x) \sum_{n=0}^{\infty} c_n \frac{x^{n+1}}{\lambda_0\lambda_1\cdots\lambda_{n+1}} + a_2(x) \sum_{n=0}^{\infty} c_n \frac{x^{n+2}}{\lambda_0\lambda_1\cdots\lambda_{n+2}} + \cdots. \tag{3.22}$$

---

[11] R. D. Carmichael, "A general expansion theorem with applications to certain integral equations of infinite order." An abstract of this paper is given in *Bulletin of the American Mathematical Society*, vol. 40 (1934), p. 211. The theorem, although actually stated for the case of positive λ's, carries with it the case of complex λ's, under the above conditions.

Either of the expansions of $F(x)$ converges absolutely and uniformly in any whatever circle interior to the circle $|x| = a$, and is valid for every interior point of the latter circle.

In the following paragraph we will use the operator B defined as in equation (3.17) where the λ's satisfy the additional conditions (3.19). By a method analogous to that used by Carmichael (*op. cit.*) we obtain theorems with reference to the operator B similar to those obtained by Carmichael with reference to the differential operator. Because of their similarity as to proof and statement we merely state one.

THEOREM 3.3. *Under the hypotheses on $F(x)$, $a_1(x)$, $a_2(x)$, $\cdots$, stated in the General Expansion Theorem, the mixed equation*

$$(3.23) \quad F(x) = B^n v(x) + a_1(x) B^{n-1} v(x) + \cdots + a_n(x) v(x) + a_{n+1}(x) B^{-1} v(x) + a_{n+2}(x) B^{-2} v(x) + \cdots,$$

where n is a positive integer, has one and just one solution $v(x)$ which is analytic at $x = 0$ and which satisfies the initial conditions $v(0) = \zeta_0$, $Bv(0) = \zeta_1, \cdots, B^{n-1}v(0) = \zeta_{n-1}$ where $\zeta_0, \zeta_1, \cdots, \zeta_{n-1}$ are any preassigned constants. This solution is analytic in the interior of the circle $|x| = a$.

UNIVERSITY OF ILLINOIS.

ON THE BOUNDARY OF THE RANGE OF VALUES OF $\zeta(s)$.

By Richard Kershner and Aurel Wintner.

In connection with his investigation of the distribution of the values of the Riemann zeta function,[1] Bohr[2] has studied the vectorial addition[3] of convex curves and has proven, among other things, that if the vectorial sum of a finite or infinite sequence of convex curves is a bounded set, then it is either a closed convex region or a ring-shaped region bounded by two non-intersecting convex curves. The method of Bohr is entirely geometrical and it is, therefore, impossible to apply it to a discussion of smoothness properties of the boundary of the vectorial sum, at least if the sum is infinite. Following a suggestion of one of the present authors, Haviland[4] applied the supporting function (Stützfunktion) of Brunn and Minkowski to the study of Bohr's problem, using the fact that the supporting function of the outer boundary of the vectorial sum is the sum of the supporting functions of the added curves. Bohr and Jessen[5] have recently shown that, with the use of the supporting functions, it is possible to determine the set of those $\sigma_0 > 1$ for which the closure of the values attained by the logarithm of the Riemann zeta function on the line $\sigma = \sigma_0$ is a convex region; for the remaining values of $\sigma_0 > 1$ this closure is ring-shaped. It has been shown by one of the present authors[6] that a rule similar to the above mentioned rule for the supporting function of the outer boundary holds for the supporting function of the inner boundary, at least on arcs of the inner boundary which are free of corners.

In the present paper the method of supporting functions will be applied to the question of regular analyticity of the boundary curves of the infinite vectorial sum in cases of the type of the logarithm of the Riemann zeta function. In particular, it will be shown that the outer boundary of the closure

[1] Cf. E. C. Titchmarsh, *The Zeta Function of Riemann*, Cambridge, 1930, Chapter IV.

[2] H. Bohr, "Om Addition af uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab*, Forhandlinger, 1913, pp. 325-366.

[3] By the vectorial sum $A_1 + A_2$ of two sets A_1 and A_2 is meant the set of all points $z = z_1 + z_2$, where $z_1 \subset A_1$ and $z_2 \subset A_2$. It is clear that this addition is associative and commutative. By an infinite vectorial sum $A_1 + A_2 + \cdots$ is meant the set of points z which may be expressed in at least one way as a limit of points $z^{(n)} = z_1 + z_2 + \cdots z_n$, where $z_k \subset A_k$.

[4] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[5] H. Bohr and B. Jessen, "On the distribution of the values of the Riemann zeta function," *American Journal of Mathematics*, vol. 58 (1936), pp. 35-44.

[6] R. Kershner, "On the addition of convex curves." To appear later. For an abstract, cf. *Bulletin of the American Mathematical Society*, vol. 42 (1936). Record of the February, 1936, meeting.

of the values attained by $\log\zeta(s)$ on a fixed line $\sigma=\sigma_0>1$ is a regular analytic curve. Due to the explicit formula for the supporting function of the inner curve, the analyticity of the inner curve may be treated similarly.[6] The result is of particular interest in view of the fact that the density of the asymptotic distribution cannot be analytic at points of the boundary. The transition from the range of $\log\zeta(\sigma_0+it)$ to that of $\zeta(\sigma_0+it)$ requires but a trivial exponential mapping.

It is known that if the power series

$$(1)\qquad p(w)=a_1w+a_2w^2+\cdots,\quad\text{where}\quad a_1\neq 0,$$

is convergent in the vicinity of $w=0$, then there exists an $R>0$ such that $p(w)$ is regular and schlicht in $|w|\leqq R$ and that[7] the curve $z=p(re^{i\theta})$, where $0\leqq\theta<2\pi$, is, for every fixed positive value of $r\leqq R$, a regular analytic convex Jordan curve in the (x,y)-plane, where $z=x+iy$. Let $r_1,r_2,\cdots$ be an infinite sequence of positive numbers such that $r_n<R$ for every n and that $r_1+r_2+\cdots$ is convergent. Let C_n denote the convex curve

$$(2)\qquad z=z_n(\theta)=p(r_ne^{i\theta})$$

and $T_n(\theta)$ the tangent to C_n at the point $z=p(r_ne^{i\theta})$. Then the equation of $T_n(\theta)$ is

$$(3)\qquad \xi\cos\phi+\eta\sin\phi-h_n(\phi)=0,$$

where

$$(4)\qquad \cos\phi=-\dot y_n(\theta)/|\dot z_n(\theta)|,\qquad \sin\phi=\dot x_n(\theta)/|\dot z_n(\theta)|$$

and

$$(5)\qquad h_n(\phi)=x_n(\theta)\cos\phi+y_n(\theta)\sin\phi,$$

it being understood that $z_n=x_n+iy_n$ and that the dot denotes differentiation with respect to θ. Since

$$e^{i\phi}=\{-\dot y_n(\theta)+i\dot x_n(\theta)\}/|\dot z_n(\theta)|=i\dot z_n(\theta)/|\dot z_n(\theta)|,$$

it is clear from (2) that

$$(6)\qquad e^{i\phi}=-e^{i\theta}p'(r_ne^{i\theta})/|p'(r_ne^{i\theta})|,$$

where the prime denotes differentiation of $p(w)$ with respect to w. On placing $\bar p(w)=\bar a_1w+\bar a_2w^2+\cdots$, it follows from (6) that

$$e^{i\phi}=-e^{i\theta}p'(r_ne^{i\theta})/\{p'(r_ne^{i\theta})\bar p'(r_ne^{-i\theta})\}^{1/2},$$

so that

$$(7)\qquad \phi=\phi_n(\theta)=\pi+\theta-i\log p'(r_ne^{i\theta}) + (i/2)\log p'(r_ne^{i\theta})+(i/2)\log\bar p'(r_ne^{-i\theta}).$$

---

[7] Cf., e. g., G. Pólya und G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1, Berlin, 1925, p. 105, no. 108.

Hence

$$(8) \qquad \frac{d\phi_n(\theta)}{d\theta} = 1 + \frac{r_n e^{i\theta} p''(r_n e^{i\theta})}{p'(r_n e^{i\theta})} - \frac{r_n e^{i\theta} p''(r_n e^{i\theta})}{2p'(r_n e^{i\theta})} + \frac{r_n e^{-i\theta} \bar{p}''(r_n e^{-i\theta})}{2\bar{p}'(r_n e^{-i\theta})}.$$

Now $p(w)$ and, consequently, $\bar{p}(w)$ are regular and schlicht in $|w| \leqq R$, which implies that the absolute values of their first derivatives have a positive lower bound β in this circle. The absolute values of the second derivatives of $p(w)$ and $\bar{p}(w)$ have in this circle a finite upper bound γ. Consequently, since $r_n \leqq R$, it is seen from (8) that

$$\left| \frac{d\phi_n(\theta)}{d\theta} \right| \geqq 1 - \alpha r_n,$$

where the positive number $\alpha = 2\gamma/\beta$ is independent of both n and θ. Thus, since $r_n \to 0$, there exists a sufficiently large n_0 such that

$$(9) \qquad \left| \frac{d\phi_n(\theta)}{d\theta} \right| \geqq \tfrac{1}{2} \quad \text{whenever} \quad n \geqq n_0.$$

Now the equation (6) may be considered as defining θ as a univalued, continuous function

$$(10) \qquad \theta = \theta_n(\phi)$$

of the normal inclination ϕ of $T_n(\theta)$. In fact, since C_n is a regular analytic convex Jordan curve, the two angular parameters θ and ϕ are in a continuous one-to-one correspondence, so that (10) is the inverse function of the function (7). Finally, it is seen from (3) and (5) that the supporting function of C_n is

$$(11) \qquad h_n(\phi) = x_n(\theta_n(\phi)) \cos \phi + y_n(\theta_n(\phi)) \sin \phi.$$

Let $\rho > 0$ be so small that the functions (2), where $n = 1, 2, \cdots$, are regular and uniformly bounded functions of the *complex* variable θ in the rectangle

$$(12) \qquad -\rho < \Re\theta < 2\pi + \rho, \qquad -\rho < \Im\theta < \rho.$$

The existence of such a $\rho > 0$, which is independent of n, is obvious from (2), since, on the one hand, every r_n is chosen less than the convergence radius of (1) and, on the other hand, $r_n \to 0$ in view of the convergence of $r_1 + r_2 + \cdots$. Since $p'(w)$ and $\bar{p}'(w)$ do not vanish in $|w| \leqq R$, one may choose n_0 so large that not only does (9) hold but also the functions $\phi = \phi_n(\theta)$, when defined in the *complex* region (12) by the explicit formula (7), are regular and in absolute value less than a number M which is independent of $n (\geqq n_0)$.

Let $\theta = \theta_0$ be a fixed real angle in the interval $0 \leqq \theta_0 \leqq 2\pi$ and let ϕ_0 be the corresponding angle ϕ. Then, in the circle $|\theta - \theta_0| < \rho$, the functions $\phi = \phi_n(\theta)$, where $\phi_0 = \phi_n(\theta_0)$, are regular analytic and in absolute value less

than M for every $n \geqq n_0$. Furthermore, $\left| \dfrac{d\phi_n(\theta)}{d\theta} \right|_{\theta=\theta_0} \geqq \frac{1}{2}$ by (9). Hence, by the proof of the local theorem on the inverse function of an analytic function, there exists [8] a positive constant τ which is independent of θ_0 and of n $(\geqq n_0)$ and is such that the inverse function $\theta = \theta_n(\phi, \theta_0, \phi_0)$, where $\phi_0 = \phi_n(\theta_0)$, is regular analytic and $| \theta - \theta_0 | < \rho$ in the circle $| \phi - \phi_0 | < \tau$. Now it is obvious from the monodromy theorem that $\theta_n(\phi, \theta_0, \phi_0)$ is independent of θ_0 and of $\phi_0 = \phi_n(\theta_0)$, so that it may be denoted simply by $\theta_n(\phi)$. Thus, if $n \geqq n_0$, all functions $\theta_n(\phi)$ are regular analytic and $\theta = \theta_n(\phi)$ is within the rectangle (12), hence the functions $\theta_{n_0}(\phi), \theta_{n_0+1}(\phi), \cdots$ are uniformly bounded, if ϕ is in the rectangle

$$(13) \qquad \mu - \tau/2 \leqq \Re\phi \leqq \mu + 2\pi + \tau/2, \qquad -\tau/2 < \Im\phi < \tau/2,$$

where $\mu = \phi_n(0)$ is the normal inclination of $T_n(0)$. Consequently, the functions $\theta_n(\phi)$, which were defined for $\mu \leqq \phi_0 \leqq \mu + 2\pi$ by (10), are regular and uniformly bounded in the rectangle (13) of the complex ϕ-plane for every $n \geqq n_0$. Finally, since the intervals $\mu \leqq \phi_0 \leqq \mu + 2\pi$ and $0 \leqq \theta_0 \leqq 2\pi$ are in a one-to-one correspondence, and since the derivative $d\phi/d\theta$ is, for real θ, distinct from zero for every n, it is clear from the definition (10) of $\theta_n(\phi)$ that there exists, for every fixed n, a rectangle R_n in the complex ϕ-plane such that R_n contains the interval $\mu \leqq \phi_0 \leqq \mu + 2\pi$ in its interior and the function (10) is regular and bounded in R_n. Consequently, there exists in the complex ϕ-plane a rectangle R_0 in which *all* functions $\theta_n(\phi)$ are regular and uniformly bounded and which contains the interval $\mu \leqq \phi_0 \leqq \mu + 2\pi$ in its interior. In fact, if Q denotes the rectangle (13), the common part of the finite number of rectangles Q; $R_1, R_2, \cdots, R_{n_0-1}$ will be such an R_0.

Now it will be shown that

$$(14) \qquad H(\phi) = \sum_{n=1}^{\infty} h_n(\phi)$$

is regular analytic in the interval $\mu \leqq \phi \leqq \mu + 2\pi$, i. e., in an open rectangle of the complex ϕ-plane which contains this interval. It has been shown above that if ϕ is in the rectangle R_0, then $| \theta_n(\phi) | < L$ for some positive constant L. On the other hand, it is clear from (1) and (2), where $z_n = x_n + iy_n$, that, since $r_n \to 0$, one may choose a sufficiently large $K = K_L$ and then a positive constant $A = A_K$ such that, if $| \theta | < L$ and $n \geqq K$, the functions $x_n(\theta)$ and $y_n(\theta)$ are regular and in absolute value not larger than Ar_n. Consequently, $x_n(\theta_n(\phi))$ and $y_n(\theta_n(\phi))$ are regular and in absolute value not larger than

---

[8] Cf. e. g., E. Landau, "Der Picard-Schottkysche Satz und die Blochsche Konstante," *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalische-mathematische Klasse*, 1926, p. 470.

Ar_n if ϕ is in the rectangle R_0 and $n \geqq K$. Hence, since $r_1 + r_2 + \cdots$ is convergent by assumption, the series

$$x(\phi) = \sum_{n=1}^{\infty} x_n(\theta_n(\phi)), \qquad y(\phi) = \sum_{n=1}^{\infty} y_n(\theta_n(\phi))$$

have regular analytic terms and are uniformly convergent in the rectangle R_0 of the ϕ-plane. It follows, therefore, from (11) that the function (14) is regular in R_0. This completes the proof of the analyticity of (14) in a rectangle which contains the interval $\mu \leqq \phi \leqq \mu + 2\pi$ in its interior.

Now $H(\phi)$ is [9] the supporting function of the outer curve. On introducing into the regular analytic functions $x = x(\phi)$, $y = y(\phi)$ the new independent variable

$$s = \int^{\phi} [(x'(\psi))^2 + (y'(\psi))^2]^{1/2}\, d\psi$$

instead of ϕ, it follows that the outer boundary of the region $C_1 + C_2 + \cdots$ is a regular analytic curve.

If $p(w)$ is the power series

$$p(w) = -\log(1 - w), \quad \text{where} \quad p(0) = 0,$$

then $p(w)$ is regular and schlicht for $|w| < 1$ and the image of the circle $|w| = r$ is convex for every $r < 1$. Now it is known [10] that the closure of the values of the logarithm of the Riemann zeta function, $\log \zeta(s)$, on a fixed line $\Re s = \sigma > 1$, is the region $C_1 + C_2 + \cdots$, where C_n is the curve

$$z = -\log(1 - p_n^{-\sigma} e^{i\theta})$$

and p_n is the n-th prime number. Since $p_n^{-\sigma} < 1$ for every n and $p_1^{-\sigma} + p_2^{-\sigma} + \cdots$ is convergent for $\sigma > 1$, it follows that the outer boundary of the closure of the values of $\log \zeta(s)$ on a fixed line $\sigma > 1$ is a regular analytic convex curve. The inner curve, if any, is similarly treated up to its possible corners.

THE JOHNS HOPKINS UNIVERSITY.

ERRATUM.

Professor H. Cramér has kindly pointed out to me that Theorem XVIII of my paper "On a class of Fourier transforms" (pp. 45-90 of vol. 58 (1936) of this JOURNAL) is false unless one reads at the end of this Theorem "a Gauss-Maxwell law" instead of "the Gauss-Maxwell law of radial symmetry." This change is necessitated by the fact that on p. 81 the statement "the matrix $\| -\frac{1}{2}\mu_{pq} \|$ is $C(e_u)$ times the unit matrix" is erroneous. That Theorem XVIII becomes correct by omitting the restriction of radial symmetry is seen from the central limit theorem.

A. WINTNER.

---

[9] Cf. E. K. Haviland, *loc. cit.*

[10] Cf., e. g., E. C. Titchmarsh, *loc. cit.*, pp. 64-67.

THE SAMPLING THEORY OF SYSTEMS OF VARIANCES, COVARIANCES AND INTRACLASS COVARIANCES.

By S. S. Wilks.

It is well known that the type of data to which the method of intraclass correlation is applicable can be effectively and accurately treated by the method of the Analysis of Variance. The connection between the two methods has been given by Fisher[1] for the case of one variable, and it is quite evident that the latter method has the greater practical value in determining, statistically, the relative importance of two groups of factors causing variation. However, the method of intraclass correlation has theoretical value in that it provides estimates of the parameters which characterize the distribution of the variates which constitute a "family." For example, under the assumption of normality, if the mean and variance of each of the variates $x_1, x_2, \cdots x_k$ is m and σ^2 and the coefficient of correlation between any pair is ρ, then the distribution of this k-variate "family" is the normal multivariate distribution characterized, however, by only the parameters m, σ^2 and ρ. ρ is called the intraclass correlation coefficient for the k variates. The method of intraclass correlation consists essentially in estimating m, σ^2 and ρ and making a significance test appropriate to the hypothesis under consideration.

In the present note we shall demonstrate a general method for handling the sampling theory of systems of variances, covariances and intraclass covariances, and illustrate the method on the generalized intraclass correlation problem.

In order to present the essentials of the method, we shall consider the following general problem. Let a system of variates $y_p = x_p - m_p$ $(p = 1, 2, \cdots n)$ be distributed according to the law

$$(1) \qquad U_n(A_{pq}, y_p) = \frac{|A_{pq}|^{1/2}}{(\pi)^{n/2}} \exp(-\Sigma A_{pq} y_p y_q) \prod_{p=1}^{n} dy_p$$

where the matrix $\| A_{pq} \|$ is positive definite with determinant $| A_{pq} |$. The characteristic function of the quantities $\xi_{pq} = y_p y_q$ is given by

$$(2) \qquad \begin{aligned} \phi = E(\exp(i\Sigma t_{pq}\xi_{pq})) &= \int \exp(i\Sigma t_{pq}\xi_{pq}) \cdot U_n(A_{pq}, y_p) \\ &= | A_{pq} |^{1/2} \, | A_{pq} - i t_{pq} |^{-(1/2)} \end{aligned}$$

[1] R. A. Fisher, *Statistical Methods for Research Workers*, 4th edition (1932), p. 205. Also see R. A. Fisher, *Metron*, vol. 1, no. 4 (1921), pp. 3-32.

where $t_{pq} = t_{qp}$ and the integral is taken over all values of the y_p. Now suppose ϕ factors into determinants of order m as follows

$$(3) \qquad \phi = \prod_{\alpha=1}^{g} | B^{\alpha}_{uv} |^{k\alpha/2} \, | \overline{B}^{\alpha}_{uv} |^{-(k\alpha/2)}$$

where the k_α are positive integers, $\| B^{\alpha}_{uv} \|$ is positive definite and

$$(4) \qquad B^{\alpha}_{uv} = \sum_{\beta=1}^{g} c_{\alpha\beta} A^{\beta}_{uv}, \qquad \overline{B}^{\alpha}_{uv} = B^{\alpha}_{uv} - i \sum_{\beta=1}^{g} c_{\alpha\beta} t^{\beta}_{uv}, \qquad | c_{\alpha\beta} | \neq 0$$

the A^{α}_{uv} (where $A^{\alpha}_{uv} = A^{\alpha}_{vu}$) being elements of the matrix $\| A_{pq} \|$. The corresponding elements in $\| t_{pq} \|$ will be denoted by t^{α}_{uv}. $\{A^{\alpha}_{uv}\}$ and $\{t^{\alpha}_{uv}\}$ will be the sets of distinct elements in $\| A_{pq} \|$ and $\| t_{pq} \|$ respectively. Therefore, since the ξ_{pq} are all distinct, we have

$$(5) \qquad \sum_{p,q} t_{pq} \xi_{pq} = \sum_{u,v,\beta} t^{\beta}_{uv} \eta^{\beta}_{uv}$$

where $\eta^{\beta}_{uv} = \Sigma \xi_{pq}$, summed over all values of p and q for which $t_{pq} = t^{\beta}_{uv}$.

Now if we let $\sum_{\beta=1}^{g} c_{\alpha\beta} t^{\beta}_{uv} = s^{\alpha}_{uv}$ we find $t^{\beta}_{uv} = \sum_{\alpha=1}^{g} C_{\alpha\beta} s^{\alpha}_{uv}$ where $\| C_{\alpha\beta} \|$ is the reciprocal of $\| c_{\alpha\beta} \|$. To find the system of quantities of which

$$(6) \qquad \phi_\alpha = | B^{\alpha}_{uv} |^{k\alpha/2} \cdot | B^{\alpha}_{uv} - i s^{\alpha}_{uv} |^{-(k\alpha/2)}$$

is the characteristic function, we express the t^{β}_{uv} in the right side of (5) in terms of the s^{β}_{uv}. Accordingly, we find that ϕ_α is the characteristic function of the quantities

$$(7) \qquad b^{\alpha}_{uv} = \sum_{\beta} C_{\alpha\beta} \eta^{\beta}_{uv}.$$

Since $\phi = \prod_{\alpha=1}^{g} \phi_\alpha$ the systems $\{b^{\alpha}_{uv}\}$ $(\alpha = 1, 2, \cdots, g)$ are independently distributed. If $k_\alpha > m + 1$, the distribution of the quantities $\{b^{\alpha}_{uv}\}$ can be shown [2] from ϕ_α to be

$$(8) \qquad V_m(B^{\alpha}_{uv}, b^{\alpha}_{uv}, k_\alpha) = \frac{| B^{\alpha}_{uv} |^{k\alpha/2} \, | b^{\alpha}_{uv} |^{(k\alpha-m-1)/2} \exp(-\sum_{u,v} B^{\alpha}_{uv} b^{\alpha}_{uv})}{\pi^{m(m-1)/4} \prod_{i=1}^{m} \Gamma^{-[(k\alpha+1-i)/2]}} \prod_{u,v} db^{\alpha}_{uv}.$$

The distributions given by (8) for all α's for which $k_\alpha > m$ are the basic

[2] J. Wishart and M. S. Bartlett, *Proceedings of the Cambridge Philosophical Society*, vol. 29 (1933), pp. 260-270.

distributions from which distributions of systems of variances, covariances and intraclass covariances can be found by transforming the $b^a{}_{uv}$. The foregoing method can be conveniently summarized in the following

THEOREM. *If the characteristic function* $|A_{pq}|^{1/2}|A_{pq}-it_{pq}|^{-(1/2)}$ $(p,q=1,2,\cdots,n;\ \|A_{pq}\|$ *positive definite) of a system of second order products* $y_py_q=\xi_{pq}$ *factors into the form*

$$\prod_{a=1}^{g}\left|\sum_{\beta}c_{a\beta}A^{\beta}{}_{uv}\right|^{k_a/2}\cdot\left|\sum_{\beta}c_{a\beta}(A^{\beta}{}_{uv}-it^{\beta}{}_{uv})\right|^{-(k_a/2)}$$

where $u,v=1,2,\cdots m$; $|c_{a\beta}|\neq 0$, *and* $\|C_{a\beta}\|=\|c_{a\beta}\|^{-1}$, *and the quantities* $A^{\beta}{}_{uv}$, $t^{\beta}{}_{uv}$ *and* $\eta^{\beta}{}_{uv}$ *are defined as in* (4) *and* (5), *then the systems* $\{b^a{}_{uv}=\sum_{\beta}C_{a\beta}\eta^{\beta}{}_{uv}\}$ *are independent of each other in the probability sense and if* $k_a>m+1$, *the distribution of the system* $\{b^a{}_{uv}\}$ *is given by* $V_m\left(\sum_{\beta}c_{a\beta}A^{\beta}{}_{uv}, b^a{}_{uv}, k_a\right)$.

Now, as an illustration of the method, let us consider an interrelated system of families $F_1, F_2,\cdots F_m$ of k variates each, where $x_{u1}, x_{u2},\cdots x_{uk}$ are the variates in F_u. Suppose the system is normally distributed such that m_u and $\sigma_u{}^2$ are the mean and variance of each variate in F_u, ρ_{uv} the correlation ($\sigma_u\sigma_v\rho_{uv}$ the covariance) between x_{ui} and x_{vi} $(i=1,2,\cdots k)$ and ρ'_{uv} the intraclass correlation ($\sigma_u\sigma_v\rho'_{uv}$ the intraclass covariance) between x_{ui} and x_{vj} $(i,j=1,2,\cdots k;\ i\neq j)$. Here, of course, $\rho_{uv}=\rho_{vu}$, $\rho'_{uv}=\rho'_{vu}$, $\rho_{uu}=1$. The elements in the symmetric and positive definite matrix Δ of correlations can be arranged so that Δ will consist of m^2 square blocks of k^2 elements each such that if θ_{uvij} is the general element in the matrix then

$$\theta_{uvij}\begin{cases}=1, & i=j,\ u=v,\\ =\rho_{uv}, & i=j,\\ =\rho'_{uv}, & i\neq j.\end{cases}\tag{9}$$

If we let the reciprocal of the matrix $\|\sigma_u\sigma_v\theta_{uvij}\|$ of variances and covariances be $\|A_{uvij}\|$ whose determinant is $|A|$, then the distribution of the x_{ui} is

$$\begin{aligned}&U_{mk}(\tfrac{1}{2}A_{uvij}, x_{ui}-m_u)\\ &\quad=\frac{|A|^{1/2}}{(2\pi)^{mk/2}}\exp\left(-\tfrac{1}{2}\Sigma A_{uvij}(x_{ui}-m_u)(x_{vj}-m_v)\right)\prod_{u,i}dx_{ui}\end{aligned}\tag{10}$$

where the sum and product are to be taken for all subscripts. It can be readily verified that A_{uvij} is of the form

$$A_{uvij} \begin{cases} = A_{uv}, & i = j \\ = A'_{uv}, & i \neq j \end{cases} \tag{11}$$

where the A_{uv} and A'_{uv} are functions of the σ's and ρ's which need not be given for our purposes. Now suppose observations have been made on N sets of families distributed according to (9). Denoting the values of the x's for the τ-th set of families by $x_{ui\tau}$ $(u = 1, 2, \cdots m;\ i = 1, 2, \cdots k)$ the probability of obtaining such a set of observed values of the x's is

$$P = \prod_{\tau=1}^{N} U_{mk}(A_{uvij}, x_{ui\tau} - m_u). \tag{12}$$

Now let us consider the product sums

$$\left.\begin{aligned} a_{uvij} &= \sum_{\tau} (x_{ui\tau} - \bar{x}_{ui})(x_{vj\tau} - \bar{x}_{vj}) \\ b_{uvij} &= N(\bar{x}_{ui} - m_u)(\bar{x}_{vj} - m_v) \end{aligned}\right\} (u, v = 1, 2 \cdots m;\ i, j = 1, 2 \cdots k) \tag{13}$$

where $\bar{x}_{ui} = (1/N) \sum_{\tau} x_{ui\tau}$. It is known from the sampling properties of normally distributed variables that the distribution of the system of means is independent of that of the second order product moments of deviations from means. The characteristic function of the a's and b's is given by the expression

$$\begin{aligned} \phi &= E[\exp(\Sigma\alpha_{uvij}a_{uvij} + \Sigma\beta_{uvij}b_{uvij})] \\ &= |\tfrac{1}{2}A|^{N/2}\, |\tfrac{1}{2}A_{uvij} - \alpha_{uvij}|^{-[(N-1)/2]} \cdot |\tfrac{1}{2}A_{uvij} - \beta_{uvij}|^{-(1/2)}. \end{aligned} \tag{14}$$

To adhere to the usual definition of a characteristic function the α's and β's will be purely imaginary numbers; all other numbers will be real. We now let

$$\begin{aligned} \alpha_{uvij} &\begin{cases} = \alpha_{uv}, & i = j \\ = \alpha'_{uv}, & i \neq j \end{cases} \\ \beta_{uvij} &\begin{cases} = \beta_{uv}, & i = j \\ = \beta'_{uv}, & i \neq j. \end{cases} \end{aligned} \tag{15}$$

Making use of (11) we find that ϕ becomes the characteristic function of the quantities in the square brackets in the following expression

$$\begin{aligned} \phi' = E\{\exp(&\sum_{u,v} \alpha_{uv}[\sum_{i} a_{uvii}] + \sum_{u,v} \alpha'_{uv}[\sum_{i\neq j} a_{uvij}] \\ &+ \sum_{u,v} \beta_{uv}[\sum_{i} b_{uvii}] + \sum_{u,v} \beta'_{uv}[\sum_{i\neq j} b_{uvij}])\}. \end{aligned} \tag{16}$$

Using equations (11) and (15) in (14) we find that ϕ' factors into the form

$$\begin{aligned} &|\tfrac{1}{2}A|^{N/2}\, |\tfrac{1}{2}C_{uv} - \gamma_{uv}|^{-[(N-1)(k-1)/2]} \cdot |\tfrac{1}{2}C_{uv} - \gamma'_{uv}|^{-[(k-1)/2]} \\ &\quad \cdot |\tfrac{1}{2}D_{uv} - \delta_{uv}|^{-[(N-1)/2]} \cdot |\tfrac{1}{2}D_{uv} - \delta'_{uv}|^{-(1/2)}, \end{aligned} \tag{17}$$

where

$$(18)\quad \begin{aligned} \gamma_{uv} &= \alpha_{uv} - \alpha'_{uv}, & \delta_{uv} &= \alpha_{uv} + (k-1)\alpha'_{uv}, & C_{uv} &= A_{uv} - A'_{uv} \\ \gamma'_{uv} &= \beta_{uv} - \beta'_{uv}, & \delta'_{uv} &= \beta_{uv} + (k-1)\beta'_{uv}, & D_{uv} &= A_{uv} + (k-1)A'_{uv}. \end{aligned}$$

In terms of the σ_u, ρ_{uv} and ρ'_{uv} the matrices $\| C_{uv} \|$ and $\| D_{uv} \|$ are the reciprocals of $\| \sigma_u\sigma_v(\rho_{uv} - \rho'_{uv}) \|$ and $\| \sigma_u\sigma_v(\rho_{uv} + (k-1)\rho'_{uv}) \|$ respectively. Let the α's and β's in (16) be transformed to the γ's and δ's by the equations (18), afterwards setting $\gamma'_{uv} = \gamma_{uv}$. ϕ' becomes

$$(19)\qquad \phi'' = E\{\exp(\sum_{u,v} [\gamma_{uv}c_{uv} + \delta_{uv}d_{uv} + \delta'_{uv}e_{uv}])\}$$

where

$$(20)\qquad \begin{aligned} c_{uv} &= \frac{k-1}{k}\sum_i (a_{uvii} + b_{uvii}) - \frac{1}{k}\sum_{i \neq j}(a_{uvij} + b_{uvij}) \\ d_{uv} &= \frac{1}{k}\sum_{i,j} a_{uvij} \\ e_{uv} &= \frac{1}{k}\sum_{i,j} b_{uvij}. \end{aligned}$$

Thus, the value of ϕ'' in (19) is given by (17) with the γ'_{uv} replaced by γ_{uv}.

Since the characteristic function of the systems $\{c_{uv}\}$ and $\{d_{uv}\}$ is the product of the characteristic functions for the two systems, each having the form of ϕ_a as given by (6), the distributions of $\{c_{uv}\}$ and $\{d_{uv}\}$ are therefore $V_m(\frac{1}{2}C_{uv}, c_{uv}, N(k-1))$ and $V_m(\frac{1}{2}D_{uv}, d_{uv}, N-1)$ respectively. If we let

$$(21)\qquad \bar{x}_u = \frac{1}{k}\sum_{i=1}^{k} \bar{x}_{ui},$$

then $e_{uv} = Nk(\bar{x}_u - m_u)(\bar{x}_v - m_v)$ and the quantities $\sqrt{Nk}(\bar{x}_u - m_u)$ are distributed according to the normal law $U_m(\frac{1}{2}D_{uv}, \sqrt{Nk}(\bar{x}_u - m_u))$ whose functional form is given by (1), and independently of the c_{uv} and d_{uv}.

Let the values of the a's and b's given by (13) be substituted in (20), and let

$$(22)\qquad \begin{aligned} s_{uv} &= \sum_{a,i} (x_{uia} - \bar{x}_u)(x_{via} - \bar{x}_v) \\ s'_{uv} &= \sum_{a, i \neq j} (x_{uia} - \bar{x}_u)(x_{vja} - \bar{x}_v). \end{aligned}$$

Then

$$(23)\qquad \begin{aligned} c_{uv} &= \left(1 - \frac{1}{k}\right) s_{uv} - \frac{1}{k} s'_{uv} \\ d_{uv} &= \frac{1}{k}(s_{uv} + s'_{uv}). \end{aligned}$$

The sample values of the variance σ_u^2, covariance $\sigma_u\sigma_v\rho_{uv}$ and intraclass covariance $\sigma_u\sigma_v\rho'_{uv}$ are given by

(24) $$s_u^2 = \frac{s_{uu}}{Nk}, \qquad \bar{s}_{uv} = \frac{s_{uv}}{Nk}, \qquad \bar{s}'_{uv} = \frac{s'_{uv}}{Nk(k-1)}.$$

The sample values of ρ_{uv} and ρ'_{uv} are respectively

(25) $$r_{uv} = \frac{s_{uv}}{\sqrt{s_{uu}s_{vv}}}, \qquad r'_{uv} = \frac{s'_{uv}}{(k-1)\sqrt{s_{uu}s_{vv}}}.$$

Incidentally, it should be remarked that $\bar{x}_u$, s_u^2, r_{uv}, and r'_{uv} are the optimum estimates of the parameters m_u, σ_u^2, ρ_{uv}, and ρ'_{uv}; that is, the values of the parameters which maximize P in (12) for a given set of x's.

The distribution of the $\bar{s}_{uv}$ and $\bar{s}'_{uv}$ will be given by the product $V_m(\frac{1}{2}C_{uv}, c_{uv}, N(k-1)) \cdot V_m(\frac{1}{2}D_{uv}, d_{uv}, N-1)$ after applying the transformations (23) and (24). The result is

(26) $$V'_m(\tfrac{1}{2}C_{uv}, N(k-1)(\bar{s}_{uv} - \bar{s}'_{uv}), N(k-1)) \cdot V'_m(\tfrac{1}{2}D_{uv}, N(\bar{s}_{uv} + \overline{k-1}\bar{s}'_{uv}), N-1)\,[N^2k(k-1)]^{m(m+1)/2} \prod d\bar{s}_{uv}d\bar{s}'_{uv}$$

where $V'_m(\ )$ denotes $V_m(\ )$ with differentials omitted. Similarly, the distribution of the s_u^2, r_{uv} and r'_{uv} can be found by applying the transformations (24) and (25) to (26).

The problem of testing from the sample, that is, the information contained in the $\bar{s}_{uv}$ and $\bar{s}'_{uv}$, that the intraclass correlations ρ'_{uv} are all zero can be solved by considering the Neyman-Pearson[3] criterion for this hypothesis. The criterion is

(27) $$\lambda_{(\rho'_{uv}=0)} = \frac{|\,\bar{s}_{uv} - \bar{s}'_{uv}\,|^{N(k-1)}\,|\,\bar{s}_{uv} + (k-1)\bar{s}'_{uv}\,|^{N}}{|\,\bar{s}_{uv}\,|^{Nk}}$$

and is the ratio of the maximum of P in (12) for variations of the parameters m_u, σ_u^2 and ρ_{uv} (with the $\rho'_{uv} = 0$) to the maximum of P for variations of all parameters. (27) is clearly a function of generalized variances. The sampling moments and the distribution of λ, under the assumption of zero intraclass correlations, can be found by a method discussed by the author elsewhere.[4] The criterion actually used in the one variable problem ($u = 1$) to test the hypothesis that the $\rho' = 0$ is

(28) $$\theta = \frac{\bar{s}_{11} + \overline{k-1}\bar{s}'_{11}}{\bar{s}_{11} - s'_{11}}.$$

The analogous criterion for m variables is

[3] J. Neyman and E. S. Pearson, *Philosophical Transactions of the Royal Society of London*, Ser. A, vol. 231 (1932), p. 295.

[4] *Annals of Mathematics*, vol. 35 (1934), p. 323.

$$\frac{|\bar{s}_{uv} - \bar{s}'_{uv}|}{|\bar{s}_{uv}|} \tag{29}$$

which is a generalization of $\frac{k}{\theta + k - 1}$. The moments and distribution of the expression in (29) can also be found by the methods just cited.

The methods used in this paper can be applied to problems of finding the general sampling distributions of systems of variances and covariances associated with Latin Squares, "equalized" random blocks, and various other lay-outs used in experimental agriculture.

PRINCETON UNIVERSITY.

TOTALLY DISCONNECTED LOCALLY COMPACT RINGS.[1]

By N. Jacobson.

Introduction.

1. Miss Taussky and I[2] have recently considered the theory of locally compact separable (l. c. s.) rings and shown that it could be reduced for the most part to the consideration of two essentially different topological types, the connected and the totally disconnected. Furthermore the connected l. c. s. rings satisfying mild algebraic conditions were shown to be hypercomplex systems with finite bases over the field of real numbers and hence their structure could be given by means of the classical results on hypercomplex numbers.

In this paper I consider the totally disconnected (t. d.) l. c. s. rings restricting myself in the main to simple rings and particularly to fields. The work is divided into two parts which are quite independent of each other. Part I is concerned with the nature of the additive abelian group of a locally-compact separable totally-disconnected (l. c. s. t. d.) simple ring $\mathfrak{S}$. $\mathfrak{S}$ may be non-commutative and non-associative. It is necessary to distinguish two cases: (1) the characteristic of $\mathfrak{S}$, $\chi(\mathfrak{S}) = p$ a prime and (2) $\chi(\mathfrak{S}) = 0$. In (1) the additive group of $\mathfrak{S}$ is a direct sum of cyclic groups of order p. (The precise meaning of this type of direct sum usually involving an infinite number of summands is given below (§ 5).) In (2) $\mathfrak{S}$ is additively a finite dimensional vector space over the field of p-adic numbers P_p. It follows that $\mathfrak{S}$ is a hypercomplex system with a finite basis relative to P_p.

Now if $\mathfrak{S}$ is associative and $\chi(\mathfrak{S}) = 0$, the above result together with the known theory of hypercomplex systems over a p-adic field gives a complete solution of the problem of determining the algebraic structure of $\mathfrak{S}$. For, by Wedderburn's theorem $\mathfrak{S}$ is a system consisting of all matrices of a fixed finite degree with coefficients in a p-adic division algebra $\mathfrak{F}$. By the results of Hasse[3] $\mathfrak{F}$ is cyclic over its centrum $\mathfrak{C}$ and the precise nature of the latter can be described.

On the other hand there is no such well-developed theory which will apply directly to solve the structure problem for the case $\chi(\mathfrak{S}) = p$. It is necessary

---

[1] Presented to the Society, April 19, 1935.

[2] "Locally compact rings," *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 106-108.

[3] H. Hasse, "Über $\mathfrak{P}$-adische Schiefkörper etc.," *Mathematische Annalen*, vol. 104 (1931), referred to below as H.

to develop new methods for the solution of this problem. This has been done in part II under the further assumption that $\mathfrak{S} = \mathfrak{F}$ is a field (not necessarily commutative). Moreover the methods apply at the same time to fields of characteristic 0 and thus afford a new treatment of p-adic division algebras. The main idea underlying this treatment is that the orders (Ordnungen) of a p-adic division algebra are compact and open (c. o.) subrings and, conversely. Since the existence of c. o. subrings of $\mathfrak{F}$ can be shown directly the procedure is as follows: We first consider the ideal theory of any c. o. subring $\mathfrak{R}$ of $\mathfrak{F}$ and then prove the existence of a unique maximal c. o. subring $\mathfrak{R}_\omega$. By means of $\mathfrak{R}_\omega$ a valuation (Bewertung) of $\mathfrak{F}$ is defined. We then use the methods of Hasse to show that $\mathfrak{F}$ is a cyclic algebra over its centrum $\mathfrak{C}$. The structure of the valued commutative field $\mathfrak{C}$ can be determined as, indeed, it has been by v. Dantzig[4] and by Hasse and Schmidt.[5]

2. For a comprehensive account of the foundations of topological algebra the reader is referred to the paper by v. Dantzig (D. l. c. in footnote 4). We recall a few of the important definitions at this point.

By a *space* $\mathfrak{S}$ we shall always mean a Hausdorff space.[6] $\mathfrak{S}$ is said to be *compact* if every infinite sequence of points in it has a limit point. $\mathfrak{S}$ is *locally compact* (l. c.) if every point has a neighborhood whose closure is compact. $\mathfrak{S}$ is *separable* (s.) if it contains a denumerable set of neighborhoods which generate all the open sets by logical addition. $\mathfrak{S}$ is *connected* if it is impossible to decompose it into a logical sum $\mathfrak{S}_1 \cup \mathfrak{S}_2$ of the open and non-intersecting sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$. $\mathfrak{S}$ is *totally disconnected* (t. d.) if for every pair of distinct points a_1, a_2 there is a decomposition of the space into a sum of non-intersecting open sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$ such that $\mathfrak{S}_i \supset a_i$. If $\mathfrak{S}$ is l. c. and t. d. then it is *zero-dimensional,* i. e., every point has arbitrarily small open and closed neighborhoods.[7]

A *locally compact separable ring* $\mathfrak{S}$ is a l. c. s. space in which there is defined two binary operations $+$ and $\cdot$ such that

1. $\mathfrak{S}$ is an abelian group under $+$.

---

[4] D. van Dantzig in *Studien over topologische Algebra,* Dissertation, Amsterdam, H. J. Paris, 1931, determines the structure of all l. c. commutative and associative fields. The foundations of topological algebra may be found in this paper (in Dutch) or in van Dantzig, "Zur topologische Algebra," *Mathematische Annalen,* vol. 107 (1933), referred to as D.

[5] H. Hasse and F. K. Schmidt, "Die Struktur diskret bewerteter Körper," *Journal f. d. reine u. angew. Math.,* vol. 170 (1933), pp. 4-63.

[6] F. Hausdorff, *Grundzüge der Mengenlehre,* 1914, p. 213.

[7] K. Menger, *Dimensionstheorie,* Teubner, 1928, p. 207.

2. $a(b+c)=ab+ac$, $(b+c)a=ba+ca$.

3. $+, -, \cdot,$ and $\div$ (when it exists) are continuous operations: If $a_\mu \to a$ (converges to a) and $b_\mu \to b$, then $a_\mu \pm b_\mu \to a \pm b$ and $a_\mu b_\mu \to ab$. If a_μ^{-1} and a^{-1} exist, then $a_\mu^{-1} \to a^{-1}$.

I. The additive group of a simple ring.

3. A l. c. s. t. d. ring $\mathfrak{S}$ is a l. c. s. t. d. abelian group under addition. It is therefore natural to begin our investigation with an account of the theory of l. c. s. t. d. abelian groups (§ 3 and § 4). It may be noted in the sequel that for our present purpose (the classification of simple rings) only a portion of the group theoretic results are needed. However, we may justify the more complete discussion on the ground that it shows up the precise extent to which the ring restrictions effect the structure of its additive group, and it should be useful for future generalizations. The main results on l. c. s. t. d. abelian groups are due to v. Dantzig (dissertation) and independently to Alexander and Cohen.[8]

In the remainder of this section $\mathfrak{S}$ will denote a l. c. s. t. d. abelian group with addition as the group operation.

THEOREM 3. 1. *$\mathfrak{S}$ is complete.*

The completeness is meant in the following sense: if $a_1, a_2, \cdots$ is a Cauchy sequence, i. e. has the property that for any given neighborhood U of 0 there exists a positive integer $N(U)$ such that $a_\mu - a_\nu \in U$ if μ and $\nu > N$, then $a_1, a_2, \cdots$ converges. To prove this property suppose U is compact, i. e., has a compact closure. Since $a_{N+1} - a_N, a_{N+2} - a_N, \cdots$ all belong to U, they have a limit b. It follows easily that $a_\mu \to a_N + b$.

THEOREM 3. 2. *If U is an open and closed compact neighborhood of* 0, *then there exists an open and closed compact subgroup $\mathfrak{G} \subset U$.*[9]

Denote the set of elements $\{-u\}$ where $u \in U$ by $-U$ and $V = U \cap (-U)$ (logical intersection). Since $-U$ and U are open and closed, so is their intersection V. We denote the complement of V in $\mathfrak{S}$ by $\mathfrak{S} \mid V$.

---

[8] J. W. Alexander and L. W. Cohen in "A classification of the homology groups of compact spaces," *Annals of Mathematics*, vol. 33 (1932), pp. 538-566, deal with groups with generators, which may be shown by virtue of Theorem 3. 2 to include the c. s. t. d. groups.

[9] This theorem is due to v. Dantzig cf. his dissertation, p. 18, and also E. R. v. Kampen, "Locally compact abelian groups," *Proceedings of the National Academy of Sciences*, vol. 20 (1934).

If $U_1 \supset U_2 \supset U_3 \supset \cdots$ is a decreasing sequence of neighborhoods of 0 whose intersection is 0 ($U_\mu \to 0$) then for μ sufficiently large $(U_\mu + V) \cap (\mathfrak{S} \mid V)$ is vacuous. Otherwise we have for every μ a $u_\mu \in U_\mu$ and $v_\mu \in V$ such that $u_\mu + v_\mu = h_\mu \in \mathfrak{S}|V$. Suppose $v_{\mu_k} \to v \in V$. Since $u_{\mu_k} \to 0$, $h_{\mu_k} \to v$ which is impossible since $\mathfrak{S}|V$ is closed.

Let $\mathfrak{G}$ denote the set of elements g of $\mathfrak{S}$ such that $g + V \subset V$. $\mathfrak{G}$ is a group $\subset V$ and by the above remark $\mathfrak{G} \supset U_\mu$ for μ sufficiently large. Hence $\mathfrak{G}$ is open. Being a group $\mathfrak{G}$ is also closed.[10]

By Theorem 3.2 there exists a sequence of compact open and closed subgroups $\mathfrak{G}_\mu \to 0$. The cosets of $\mathfrak{G}_\mu$ constitute a fundamental set of neighborhoods for $\mathfrak{S}$. In the rest of I we will therefore mean by a neighborhood of 0 a compact open and closed group neighborhood.

COROLLARY. *$\sum_{\mu=1}^{\infty} a_\mu$ exists if and only if $a_\mu \to 0$.*

Set $\sum_{\mu=1}^{n} a_\mu = s_n$. By the completeness of $\mathfrak{S}$, $\sum_{\mu=1}^{\infty} a_\mu$ exists if and only if $\{s_n\}$ is a Cauchy sequence. Let U be a (compact group) neighborhood of 0. If $\mu > N(U)$, $a_\mu \in U$ and hence also for $n, m > N$

$$s_n - s_m = \begin{cases} a_{m+1} + a_{m+2} + \cdots + a_n & \text{if} \quad n > m \\ 0 & \text{if} \quad n = m \\ a_{n+1} + a_{n+2} + \cdots + a_m & \text{if} \quad n < m \end{cases}$$

is contained in U. Thus $\{s_n\}$ is a Cauchy sequence and Σa_μ exists. The converse is obvious.

4. A group $\mathfrak{S}$ is called a *p-group* (p a prime) if for every $a \in \mathfrak{S}$, $p^\mu a \to 0$. Let $\mathfrak{G}$ be a compact t. d. group and

$$\mathfrak{G}^{(p)} = \underset{\mu}{I}(p^\mu \mathfrak{G}) \equiv \mathfrak{G} \cap p\mathfrak{G} \cap p^2\mathfrak{G} \cap \cdots.$$

LEMMA 4.1. *A necessary and sufficient condition that an element $h \in \mathfrak{G}^{(p)}$ is that there exist a sequence of integers $\mu_i \to \infty$ such that $p^{\mu_i} h \to h$.*

(1) Suppose $p^{\mu_i} h \to h$. Since $p^{\mu_i} h \in p^\mu \mathfrak{G}$ for $\mu_i \geqq \mu$ and $p^\mu \mathfrak{G}$ is closed, we have $p^\mu \mathfrak{G} \supset h$. Hence $h \in \underset{\mu}{I} p^\mu \mathfrak{G} = \mathfrak{G}^{(p)}$.

(2) Let $h \in \mathfrak{G}^{(p)}$ so that $h = p^\nu g_\nu$ $(\nu = 1, 2, \cdots)$. If $g_{\nu_k} \to g$ then $h = \lim_{k \to \infty} p^{\nu_k} g$. For if k is sufficiently large $(g_{\nu_k} - g) \in U$ where U is an arbitrary (group) neighborhood of 0. Hence

[10] D. p. 609. That $\mathfrak{G}$ is closed may also be seen directly.

$$p^{\nu_k}(g_{\nu_k} - g) = (h - p^{\nu_k} g) \in U$$

or $\lim_{k \to \infty} p^{\nu_k} g = h$. Set $\mu_k = \nu_{k+1} - \nu_k$. By dropping enough terms of the sequence $\{\nu_k\}$ we may suppose that $\mu_k \to \infty$. We assert that $p^{\mu_k} h \to h$. If k is large enough, $(h - p^{\nu_k} g) \in U$ and hence

$$p^{\mu_k}(h - p^{\nu_k} g) = (p^{\mu_k} h - p^{\nu_{k+1}} g) \in U$$
$$(h - p^{\mu_k} h) = (h - p^{\nu_{k+1}} g) - (p^{\mu_k} h - p^{\nu_{k+1}} g) \in U$$

i. e. $p^{\mu_k} h \to h$.

LEMMA 4. 2. *If $p^{\mu_k} g \to 0$ then $p^{\mu} g \to 0$ $(\mu = 1, 2, \cdots)$.*

If k is large enough so that $p^{\mu_k} g \in U$ then $p^{\mu} g \in U$ also for $\mu \geqq \mu_k$.

According to this lemma we have that the elements g for which the sequence $\{p^{\mu} g\}$ has 0 for limit point form a closed subgroup $\mathfrak{G}_p$.

LEMMA 4. 3. $\mathfrak{G} = \mathfrak{G}^{(p)} \oplus \mathfrak{G}_p$.

Let g be an arbitrary element of $\mathfrak{G}$ and $p^{\nu_k} g \to h$. Since $p^{\mu_k} h \to h$ where $\mu_k = \nu_{k+1} - \nu_k$, $h \in \mathfrak{G}^{(p)}$. Suppose $p^{\mu'_k} g \to h' \in \mathfrak{G}^{(p)}$. Then $p^{\nu'_k} h' \to h$. For if U is a neighborhood of 0 and $k > K_1(U)$, $(h' - p^{\mu'_k} g) \in U$ and hence also $(p^{\nu'_k} h' - p^{\nu'_{k+1}} g) \in U$. If $k > K_2(U)$ we have besides $(p^{\nu'_{k+1}} g - h) \in U$. Hence

$$(h - p^{\nu'_k} h') \in U, \quad \text{or} \quad p^{\nu'_k} h' \to h.$$

Thus $p^{\nu'_k}(g - h') \to 0$ and by Lemma 4. 2 $p^{\nu}(g - h') \to 0$.

Set $g = h' + (g - h')$. $h' \in \mathfrak{G}^{(p)}$ and $(g - h') \in \mathfrak{G}_p$. Evidently this decomposition is unique, i. e., $\mathfrak{G}^{(p)} \cap \mathfrak{G}_p = 0$ and so $\mathfrak{G} = \mathfrak{G}^{(p)} \oplus \mathfrak{G}_p$. (This shows that h' determined above is unique and in particular $p^{\mu_k} g \to h'$ not merely $p^{\mu'_k} g$).

LEMMA 4. 4. *There exist primes p such that $\mathfrak{G}^{(p)} \neq \mathfrak{G}$ and hence such that $\mathfrak{G}_p \neq 0$.*

If $\mathfrak{H}_0$ is an open subgroup of $\mathfrak{G}$ then $\mathfrak{G} - \mathfrak{H}_0$ is compact and discrete and hence it is finite. Let p be a divisor of its order. $\mathfrak{G} - \mathfrak{H}_0$ has a subgroup $\mathfrak{H} - \mathfrak{H}_0$ of index p. Hence $\mathfrak{G} - \mathfrak{H} \cong (\mathfrak{G} - \mathfrak{H}_0) - (\mathfrak{H} - \mathfrak{H}_0)$ has order p and $p\mathfrak{G} \subset \mathfrak{H} < \mathfrak{G}$ ($<$ means properly contained in). It follows from Lemma 4. 3 that $\mathfrak{G}_p \neq 0$.

THEOREM 4. 1. *$\mathfrak{G}$ is a direct sum of p-groups.*[11]

[11] Alexander and Cohen, *loc. cit.*,[8] p. 557.

Let p_1 be a prime for which $\mathfrak{G}_{p_1} \neq 0$. We have $\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}$. Since $\mathfrak{G}^{(p_1)}$ is closed, it is compact and may be treated as $\mathfrak{G}$, i. e. $\mathfrak{G}^{(p_1)} = \mathfrak{G}^{(p_1)}_{p_2} \oplus \mathfrak{G}^{(p_1p_2)}$. Continuing this process we obtain

$$\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}_{p_2} \oplus \cdots \oplus \mathfrak{G}^{(p_1 \cdots p_{n-1})}_{p_n} \oplus \mathfrak{G}^{(p_1 \cdots p_n)}$$

where $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}} \neq 0$ for $i < n$ if $\mathfrak{G}^{(p_1 \cdots p_n)} \neq 0$. We must show that this process is uniformly convergent in the sense that $\mathfrak{G}^{(p_1 \cdots p_n)} \subset U$ where U is any neighborhood of 0 and $n > N(U)$. Choose N so that $\{p_1, p_2, \cdots, p_n\}$ for $n > N$ includes all the distinct prime factors of the order m of $\mathfrak{G} - U$. Let $h \in \mathfrak{G}^{(p_1 \cdots p_n)}$. Then there exist sequences of integers $\{\mu_k\}, \{\nu_k\}, \cdots, \{\rho_k\}$ such that $p_1^{\mu_k}h \to h$, $p_2^{\nu_k}h \to h, \cdots, p_n^{\rho_k}h \to h$. It follows that $p_1^{\mu_k}p_2^{\nu_k} \cdots p_n^{\rho_k}h \to h$. If $\mu_k, \nu_k, \cdots, \rho_k$ are sufficiently large $p_1^{\mu_k}p_2^{\nu_k} \cdots p_n^{\rho_k}h = p_1^{\mu'_k}p_2^{\nu'_k} \cdots p_n^{\rho'_k}mh \in m\mathfrak{G}$ and hence $h \in m\mathfrak{G} \subset U$ or $\mathfrak{G}^{(p_1 \cdots p_n)} \subset U$. Hence we may write

$$\mathfrak{G} = \mathfrak{G}_{p_1} \oplus \mathfrak{G}^{(p_1)}_{p_2} \oplus \mathfrak{G}^{(p_1p_2)}_{p_3} \oplus \cdots.$$

The components $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}}$ are *characteristic subgroups* of $\mathfrak{G}$, i. e. they are carried into themselves by every automorphisms of $\mathfrak{G}$.

5. Let $\mathfrak{S}$ be a l. c. s. t. d. simple[12] ring (not necessarily associative or commutative). Since $\mathfrak{S}$ is homogeneous it is either discrete or dense in itself. The first case is, of course, uninteresting for the present considerations and hence we restrict ourselves to the second.

THEOREM 5. 1. *$\mathfrak{S}$ is a p-group.*

Let $\mathfrak{S}_p$ denote the subset of elements a of $\mathfrak{S}$ such that $p^\nu a \to 0$. $\mathfrak{S}_p$ is an ideal[13] and hence either $\mathfrak{S}_p = \mathfrak{S}$ or $\mathfrak{S}_p = 0$. $\mathfrak{S}$ contains a compact and open subgroup $\mathfrak{G}$. Since $\mathfrak{S}$ is not discrete $\mathfrak{G} \neq 0$. By Lemma 4. 4 $\mathfrak{G}$ contains elements $g \neq 0$ such that $p^\nu g \to 0$ for some p. Thus $\mathfrak{S}_p \neq 0$ and so $\mathfrak{S}_p = \mathfrak{S}$.

p is unique. For, suppose that $p_1^\nu a \to 0$ and $p_2^\nu a \to 0$ where $(p_1, p_2) = 1$. Integers α_ν and β_ν can be determined so that $\alpha_\nu p_1^\nu + \beta_\nu p_2^\nu = 1$ $(\nu = 1, 2, 3, \cdots)$. If U is an arbitrary neighborhood of 0, there exists an $N(U)$ such that $p_1^\nu a$ and $p_2^\nu a \in U$ for all $\nu \geqq N$. Then $a = \alpha_\nu p_1^\nu a + \beta_\nu p_2^\nu a \in U$. Since U is arbitrary a must be 0.

We now distinguish two cases:

(1) *Characteristic p.* $p\mathfrak{S} = 0$.

Let $\mathfrak{G}_1 > \mathfrak{G}_2 > \mathfrak{G}_3 > \cdots$ be a sequence of compact and open subgroups of $\mathfrak{S}$ whose intersection is 0. $\mathfrak{G}_\mu - \mathfrak{G}_{\mu+1}$ has order a power of p and hence

[12] A ring $\mathfrak{S}$ is said to be *simple* if (0) and (1) $\equiv \mathfrak{S}$ are its only (two-sided) ideals.

[13] The term *ideal* will refer exclusively to two-sided ideals.

by inserting a finite number of compact and open subgroups between $\mathfrak{G}_\mu$ and $\mathfrak{G}_{\mu+1}$ and changing the notation, we may suppose that $\mathfrak{G}_\mu - \mathfrak{G}_{\mu+1}$ has order p. If x_μ is an element of $\mathfrak{G}_\mu | \mathfrak{G}_{\mu+1}$ then $\mathfrak{G}_\mu = (x^\mu) \oplus \mathfrak{G}_{\mu+1}$ where (x_μ) denotes the group of order p generated by x_μ. Hence $\mathfrak{G}_1 = (x_1) \oplus (x_2) \oplus \cdots \oplus x_{n-1} \oplus \mathfrak{G}_n$ and since $\mathfrak{G}_n \to 0$

$$\mathfrak{G}_1 = (x_1) \oplus (x_2) \oplus (x_3) \oplus \cdots.$$

Suppose $y_1, y_2, y_3, \cdots$ is a denumerable dense set in $\mathfrak{S}$. Then $\mathfrak{S} = (\mathfrak{G}_1 + y_1) \cup (\mathfrak{G}_1 + y_2) \cup \cdots$.[14] If y_{μ_1} is the first $y \not\in \mathfrak{G}_1$ set $y_{\mu_1} = x_{-1}$ and form $\mathfrak{G}_{-1} = (x_{-1}) \oplus \mathfrak{G}_1$. Again if y_{μ_2} is the first $y \not\in \mathfrak{G}_{-1}$ set $y_{\mu_2} = x_{-2}$ and form $\mathfrak{G}_{-2} = (x_{-2}) \oplus \mathfrak{G}_{-1}$. Continuing in this way we obtain the theorem:

THEOREM 5.2. *$\mathfrak{S}$ has a denumerable set of generators $x_1, x_2, \cdots$ and $x_{-1}, x_{-2}, \cdots$ such that every element is expressible uniquely as an infinite linear combination of the x's with coefficients mod p and in which only a finite number of the $x_{-\mu}$'s occur. Every such expression is an element of $\mathfrak{S}$.*

(2) *Characteristic* 0. $p\mathfrak{S} \neq 0$.

LEMMA 5.1. *$\mathfrak{S}$ is a hypercomplex system over the field of rational numbers* P.[15]

If m is a positive integer, then $m\mathfrak{S}$ is an ideal and hence is either $= \mathfrak{S}$ or $= 0$. But $m\mathfrak{S} = 0$ is impossible: If $m = kl$, $m\mathfrak{S} = k(l\mathfrak{S}) = 0$ and hence either $k\mathfrak{S} = 0$ or $l\mathfrak{S} = 0$. We obtain in this way that a prime q exists for which $q\mathfrak{S} = 0$. Since p is the only prime such that $p^\nu a \to 0$ for every a in $\mathfrak{S}$, $p = q$ and we have a contradiction of the assumption $p\mathfrak{S} \neq 0$.

The elements of order m form an ideal $\neq \mathfrak{S}$. Hence this ideal is $= (0)$, i. e., every element of $\mathfrak{S}$ has infinite order. From $m\mathfrak{S} = \mathfrak{S}$ follows that for every a there exists a unique a' such that $ma' = a$. We denote a' by $(1/m)a$ and define $(n/m)a = na'$ where n is any integer. If $n/m = n_1/m_1$, then $(n/m)a = (n_1/m_1)a$. Using the fact that $\mathfrak{S}$ has no elements of finite order we can easily verify the following rules:

$$(*) \qquad \begin{gathered} \alpha(a+b) = \alpha a + \alpha b \\ (\alpha+\beta)a = \alpha a + \beta a \qquad (\alpha\beta)a = \alpha(\beta a) \\ \alpha(ab) = (\alpha a)b = a(\alpha b) \end{gathered}$$

where α and β are rational numbers.

[14] This is the only point in the entire discussion at which the assumption of separability seems to be necessary. In the rest of the paper it may be replaced by the weaker Hausdorff first denumerability axiom.

[15] $\mathfrak{S}$ does not have a finite basis over P, however.

Let $\mathfrak{G}$ be a compact and open subgroup of $\mathfrak{S}$. Then $\mathfrak{G} > p\mathfrak{G} > p^2\mathfrak{G} > \cdots$ and $\underset{\nu}{I}(p^\nu\mathfrak{G}) = 0$. For otherwise by Lemma 4.1 $\mathfrak{G}$ would contain an element $h \neq 0$ for which the sequence $\{p^\nu h\}$ has h for a limit point. Let $p^{-\mu}\mathfrak{G}$ denote the compact and open subgroup of p^μ-th parts of the elements of $\mathfrak{G}$ and $\Gamma_\mu = p^\mu\mathfrak{G} | p^{\mu+1}\mathfrak{G}$ $(\mu = 0, \pm 1, \pm 2, \cdots)$. Then $\mathfrak{G} < p^{-1}\mathfrak{G} < p^{-2}\mathfrak{G} < \cdots$. Since $p^\nu a \to 0$ for every a, there exists an integer N such that $p^N a \in \mathfrak{G}$ and hence $a \in p^{-N}\mathfrak{G}$. Thus $\mathfrak{G} \cup p^{-1}\mathfrak{G} \cup p^{-2}\mathfrak{G} \cup \cdots = \mathfrak{S}$.

If $a \in \Gamma_\mu$, evidently $p^\nu a \in \Gamma_{\mu+\nu}$ for any integral ν. If α is a rational number prime to p (numerator and denominator in the reduced form of α are prime to p), then $\alpha a \in \Gamma_\mu$. For, if α is an integer prime to p, $\alpha a \in \Gamma_{\mu+\epsilon}$ where $\epsilon \geqq 0$. But $\epsilon > 0$ is impossible since in this case $\alpha a \in p^{\mu+1}\mathfrak{G}$ contradicting the fact that the order of the coset of a in $p^\mu\mathfrak{G} - p^{\mu+1}\mathfrak{G}$ is p. The result for reciprocals of integers follows from symmetry and hence it follows for all rational α's by combining the two special cases.

If $a \in \Gamma_\mu$ we denote the real-valued function $p^{-\mu}$ of a by $\| a \|$. We have then:

$$\| a + b \| \leqq \max (\| a \|, \| b \|)$$
$$\| p^\nu a \| = p^{-\nu} \| a \|$$
$$\| \alpha a \| = \| a \| \qquad \text{if} \quad (\alpha, p) = 1.$$

If we set dist $(a, b) = \| a - b \|$, we obtain a metric which gives the topology of $\mathfrak{S}$.

Lemma 5.2. *If $\{\alpha_\nu\}$ is a sequence of rational numbers, $\{\alpha_\nu a\}$ $(a \neq 0)$ converges if and only if $\{\alpha_\nu\}$ converges in the p-adic topology.*[16]

Because of the completeness of $\mathfrak{S}$ and of the p-adic field P_p it is necessary to show only that $\alpha_\nu a \to 0$ if and only if $\alpha_\nu \xrightarrow[p]{} 0$ (in the p-adic topology). This is evident from the above considerations.

Lemma 5.3. *$\mathfrak{S}$ is a hypercomplex system over the field of p-adic numbers.*

If α is a p-adic number, say $\alpha = \lim\limits_{p} \alpha_\nu$ where the α_ν are rational numbers, then we define $\alpha a = \lim \alpha_\nu a$. This is independent of the particular sequence $\{\alpha_\nu\}$ approaching α. From the continuity of addition and multiplication we obtain the validity of equations (*) for all p-adic α, β. Finally, if $\{\alpha_\nu\}$ is a sequence of p-adic numbers converging to α and $a_\nu \to a$ then $\alpha_\nu a_\nu \to \alpha a$, i. e. αa is a continuous function of α and a.

[16] For a definition of the p-adic field P_p and the p-adic topology of P see v. d. Waerden, *Moderne Algebra* I, Springer, 1930, pp. 218-220.

We denote the p-adic vector space determined by the elements $a_1, a_2, \cdots, a_r$ of $\mathfrak{S}$ by $(a_1, a_2, \cdots, a_r)$.

LEMMA 5.4. *$(a_1, a_2, \cdots, a_r)$ is closed.*

We may suppose that $a_1, \cdots, a_r$ are linearly independent (with p-adic coefficients). Let $b_\nu = \beta_{\nu 1}a_1 + \cdots + \beta_{\nu r}a_r \to b$. We must show that $b \in (a_1, a_2, \cdots, a_r)$. For each ν choose $\lambda(\nu) = 1, 2, \cdots, r$ such that $\| \beta_{\nu,\lambda(\nu)} a_{\lambda(\nu)} \| \geqq \| \beta_{\nu,\mu} a_\mu \|$ for $\mu \neq \lambda$. Since $\lambda(\nu)$ has only a finite range one of its values, say $\lambda(\nu) = 1$ occurs infinitely often. By restricting ourselves to a subsequence we may suppose that $\| \beta_{\nu 1} a_1 \| \geqq \| \beta_{\nu\mu} a_\mu \|$ for all ν and μ. Then $\| (\| \beta_{\nu 1} a_1 \|) \beta_{\nu\mu} a_\mu \| \leqq 1$ and hence we may suppose that $\lim_{\nu \to \infty} \| \beta_{\nu 1} a_1 \| \beta_{\nu\mu} a_\mu$ exists and $= \gamma_\mu a_\mu$. Since $\| (\| \beta_{\nu 1} a_1 \|) \beta_{\nu 1} a_1 \| = 1, \gamma_1 \neq 0$. If $\lim_p \| \beta_{\nu 1} a_1 \| = \infty$, then $\| \beta_{\nu 1} a_1 \| b \to 0$ and hence $\gamma_1 a_1 + \gamma_2 a_2 + \cdots + \gamma_r a_r = 0$ contrary to the linear independence of the a's. Thus $\lim_p \| \beta_{\nu 1} a_1 \|$ exists and since $\| \beta_{\nu 1} a_1 \| \geqq \| \beta_{\nu\mu} a_\mu \|$ we may suppose that $\beta_{\nu\mu} a_\mu \to \beta_\mu a_\mu$. Then

$$b = \beta_1 a_1 + \beta_2 a_2 + \cdots + \beta_r a_r \in (a_1, a_2, \cdots, a_r).$$

THEOREM 5.3. *$\mathfrak{S}$ is a hypercomplex system with a finite basis over the field of p-adic numbers.*

Choose a_1 in Γ_0. If $\mathfrak{S} \neq (a_1)$, then there exists an $a_2 \in \Gamma_0$ such that $a_2 \not\in (a_1) + p\mathfrak{G}$. For if every b in Γ_0, $\in (a_1) + p\mathfrak{G}$ then every b_ν in $\Gamma_\nu \in (a_1) + p^{\nu+1}\mathfrak{G}$. Thus

$$\begin{array}{lll} b \ = \alpha_1 a_1 + b_1 & b_1 \in \Gamma_{m_1} & m_1 > 0 \\ b_1 = \alpha_2 a_1 + b_2 & b_2 \in \Gamma_{m_2} & m_2 > m_1 \\ \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot & \cdot \quad \cdot \quad \cdot \\ b_\nu = \alpha_\nu a_1 + b_{\nu+1} & b_{\nu+1} \in \Gamma_{m_{\nu+1}} & m_{\nu+1} > m_\nu \end{array}$$

and hence $b = (\alpha_1 + \cdots + \alpha_\nu)a_1 + b_{\nu+1}$. Since $\lim b_\nu = 0$ and (a_1) is closed, $b \in (a_1)$ contradicting $(a_1) \neq \mathfrak{S}_1$. Now choose $a_2 \begin{cases} \in \Gamma_0 \\ \not\in (a_1) + p\mathfrak{G} \end{cases}$. The subgroup generated by a_1 and a_2 mod $p\mathfrak{G}$ then has order p^2. If $\mathfrak{S} \neq (a_1, a_2)$, we repeat the process and obtain an $a_3 \in \Gamma_0$ such that the order of the group generated by a_1, a_2, a_3 mod $p\mathfrak{G}$ is p^3 and so on. Since $\mathfrak{G} - p\mathfrak{G}$ is finite this process breaks off and we obtain a finite basis for $\mathfrak{S}$.

Before leaving this part we note that the above methods are applicable to cases other than the one we have considered. We shall not attempt to give

the maximum generality of this group-theoretic method but mention certain extensions which are readily made.

A compact t. d. ring $\mathfrak{G}$ *is a direct sum of p-rings.*

This follows at once from the remark that the $\mathfrak{G}^{(p_1 \cdots p_i)}_{p_{i+1}}$ in Theorem 4. 1 are characteristic subgroups. For, then they are ideals and the sum is direct in the sense of rings as well as of groups.

Theorem 5. 3 may be stated under the more general assumptions that $\mathfrak{S}$ is a hypercomplex system over P and is a p-group.

II. The structure of associative fields.

6. Let $\mathfrak{F}$ be a l. c. s. t. d. associative field (not necessarily commutative). For the present we may drop the distinction between the two types characteristic 0 and characteristic p. We assume, of course, that $\mathfrak{F}$ is not discrete.

There exists a $y \neq 0$ sufficiently near 0 such that $y\bar{U} < U$ where U is a compact neighborhood of 0 and $\bar{U}$ is its closure. Hence $U > y\bar{U} > y^2\bar{U} > \cdots$. Suppose $y^{\nu_k} \to z \neq 0$. Then $y^{\mu_k} \to 1$ where $\mu_k = \nu_{k+1} - \nu_k$. If u is any element of U, $y^{\mu}u \in y^{\mu_k}\bar{U}$ for $\mu \geqq \mu_k$ and hence $u = \lim_{i\to\infty} y^{\mu_i}u$ is contained in the closed set $y^{\mu_k}\bar{U}$. Thus $U \subset y^{\mu_k}\bar{U}$ which is impossible. Hence $y^{\nu} \to 0$.

We require the existence of such a y in proving

LEMMA 6. 1. *If* $a_\nu \to \infty$ *then* $a_\nu^{-1} \to 0$.[17]

Evidently $\{a_\nu^{-1}\}$ can have no limit point other than 0 and hence it is sufficient to show that 0 is really a limit point of this sequence. Let $y \neq 0$ be an element such that $y^\nu \to 0$ and U a compact neighborhood of 0. If infinitely many $a_\nu y^j$ for $\nu = 1, 2, \cdots$ and j fixed were contained in U, then the corresponding a_ν would be contained in the compact set $\bar{U}y^{-j}$ and hence would have a limit point. It follows that a sequence $\{b_\nu\}$ may be extracted from $\{a_\nu\}$ so that $b_\nu y^\nu \in \mathfrak{F}|U$. Since $y^\mu \to 0$ we may determine for each ν an integer k_ν such that $b_\nu y^{k_\nu} \in \mathfrak{F}|U$ but $b_\nu y^{k_\nu+1} \in U$. Evidently $k_\nu \ (\geqq \nu) \to \infty$. Because of the compactness of U we may suppose that $b_\nu y^{k_\nu+1} \to z$. Then $b_\nu y^{k_\nu} \to zy^{-1} = w \in \mathfrak{F}|U$ and hence is $\neq 0$. From $y^{-k_\nu}b_\nu^{-1} \to w^{-1}$ and $y^{k_\nu} \to 0$ follows $b_\nu^{-1} \to 0$.

Let $\mathfrak{G}$ be a compact and open subgroup of $\mathfrak{F}$ and $\mathfrak{R}(\mathfrak{G})$ the set of elements a of $\mathfrak{F}$ such that $a\mathfrak{G} \subset \mathfrak{G}$.

---

[17] $a_\nu \to \infty$ means that $\{a_\nu\}$ is absolutely divergent, i. e., has no convergent subsequence. Lemma 6. 1 is the "Perfektisierungsaxiom" for fields of van Dantzig (D, p. 612). The present proof holds for all locally compact fields satisfying the first denumerability axiom.

LEMMA 6.2. *$\mathfrak{R}(\mathfrak{G})$ is a compact and open domain of integrity (d. o. i.).*

That $\mathfrak{R}(\mathfrak{G})$ is a ring may be verified directly. Since $\mathfrak{F}$ is a field $\mathfrak{R}$ has no zero-divisors. By the continuity of multiplication we have that $\mathfrak{R}$ contains a neighborhood of 0 and hence is open. If $\{a_\nu\}$ is an infinite sequence of elements of $\mathfrak{R}$ and $g \neq 0$ belongs to $\mathfrak{G}$ then $\{a_\nu g\}$ is a sequence in $\mathfrak{G}$ and hence has a limit point h. Thus $\{a_\nu\}$ has hg^{-1} as limit point and $\mathfrak{R}$ is compact.

7. In this section we consider the ideal theory of a compact and open (c. o.) d. o. i. $\mathfrak{R}$ contained in $\mathfrak{F}$. The existence of such subrings has just been shown.

THEOREM 7.1. *$\mathfrak{R}$ satisfies the chain conditions for left-(right-) ideals.*[18]

If $\mathfrak{J}$ is a left-(right-) ideal then $\mathfrak{J} \supset \mathfrak{R}b$ where $b \in \mathfrak{J}$. If $b \neq 0$, $\mathfrak{R}b$ is open and hence $\mathfrak{J} = \sum_{b \neq 0} \mathfrak{R}b$ is open. Let $\mathfrak{J}_1 < \mathfrak{J}_2 < \mathfrak{J}_3 < \cdots$ be an increasing sequence of ideals and $\nu_1, \nu_2, \nu_3, \cdots$ the orders of the finite quotient groups $\mathfrak{R} - \mathfrak{J}_1, \mathfrak{R} - \mathfrak{J}_2, \mathfrak{R} - \mathfrak{J}_3, \cdots$. Then $\nu_1 > \nu_2 > \nu_3 > \cdots$. Hence the sequence of $\mathfrak{J}$'s is finite. Similarly, every decreasing sequence of left-(right-) ideals $\mathfrak{J}_1 > \mathfrak{J}_2 > \mathfrak{J}_3 > \cdots > \mathfrak{J}_\omega$ containing a fixed ideal $\mathfrak{J}_\omega$ is finite.

Let $\mathfrak{P}$ be the totality of elements b of $\mathfrak{R}$ for which $b\mathfrak{R} < \mathfrak{R}$. We propose to show that $\mathfrak{P}$ is a prime ideal in $\mathfrak{R}$ and that $\mathfrak{P} > \mathfrak{P}^2 > \cdots \to 0$.

LEMMA 7.1. *$\mathfrak{R}|\mathfrak{P}$ contains arbitrarily small ideals.*

If U is any neighborhood of 0, there exists a $z \neq 0$ sufficiently near 0 such that $\mathfrak{R}z\mathfrak{R} \subset U$. $\mathfrak{R}z\mathfrak{R}$ is evidently an ideal.

LEMMA 7.2. *$\mathfrak{R}|\mathfrak{P}$ is a compact group relative to multiplication.*

$\mathfrak{R}|\mathfrak{P}$ consists of those elements $a, a', \cdots$ of $\mathfrak{F}$ which satisfy $a\mathfrak{R} = \mathfrak{R}$, $a'\mathfrak{R} = \mathfrak{R}, \cdots$. It follows then that $a'a\mathfrak{R} = \mathfrak{R}$, $a^{-1}\mathfrak{R} = \mathfrak{R}$, i. e., $aa', a^{-1} \in \mathfrak{R}|\mathfrak{P}$. Hence $\mathfrak{R}|\mathfrak{P}$ is a group. Since $\mathfrak{R}|\mathfrak{P}$ is closed and contained in $\mathfrak{R}$ it is compact.

LEMMA 7.3. *An element b of $\mathfrak{R}$ belongs to $\mathfrak{P}$ if and only if $b^\nu \to 0$.*

Since b has an inverse and $\mathfrak{R} > b\mathfrak{R}$ we have $\mathfrak{R} > b\mathfrak{R} > b^2\mathfrak{R} > \cdots$. If $b^{\nu_k} \to c \neq 0$, then $b^{\mu_k} \to 1$ where $\mu_k = \nu_{k+1} - \nu_k$. But this is impossible for the reasons given in § 6. The converse is obvious.

[18] The conditions referred to are the "Teilerkettensatz" and the "eingeschrankte Vielfachenkettensatz" of E. Noether, cf. "Abstrakter Aufbaue der Idealtheorie in algebraischen Zahl-und Funktionsenkorper," *Mathematische Annalen*, vol. 96 (1926), p. 26.

If $\mathfrak{P}_1$ is the set of b's such that $\mathfrak{R}b < \mathfrak{R}$, we may show as in Lemma 7.3 that $\mathfrak{P}_1$ consists of those elements b of $\mathfrak{R}$ whose power sequence $b^\nu \to 0$. Hence $\mathfrak{P}_1 = \mathfrak{P}$.

LEMMA 7.4. *If $\mathfrak{A}$ is a finite ring, its radical $\mathfrak{N}$ consists of the properly nilpotent elements,* i. e., the elements z such that za and az are nilpotent for all a in $\mathfrak{A}$.

This lemma is well known for rings with a finite basis over a field and it may be proved in exactly the same manner for our case of finite rings.[19]

THEOREM 7.2. *$\mathfrak{P}$ is a prime ideal in $\mathfrak{R}$ and $\underset{\nu}{I}\mathfrak{P}^\nu = 0$.*[20]

If $b \in \mathfrak{P}$ and $a \in \mathfrak{R}$ then ab and $ba \in \mathfrak{P}$. Hence $\mathfrak{P}$ is invariant. Since $\mathfrak{P}$ contains a sufficiently small neighborhood of 0, it contains an ideal $\mathfrak{J}$ of $\mathfrak{R}$. Consider the finite ring $\mathfrak{A} = \mathfrak{R} - \mathfrak{J}$. In the isomorphism $\mathfrak{R} \sim \mathfrak{A}$ the elements of $\mathfrak{P}$ are precisely those elements of $\mathfrak{R}$ which correspond to properly nilpotent elements of $\mathfrak{A}$. Hence $\mathfrak{P}$ corresponds to the radical $\mathfrak{N}$ of $\mathfrak{A}$. It follows that $\mathfrak{P}$ is an ideal and $\mathfrak{P}^n \equiv 0(\mathfrak{J})$ for sufficiently high n. Since $\mathfrak{J}$ is arbitrarily small, $\underset{\nu}{I}\mathfrak{P}^\nu = 0$. $\mathfrak{P}$ is evidently prime.

From this theorem and Wedderburn's theorem on finite fields [21] we obtain

THEOREM 7.3. *$\mathfrak{R} - \mathfrak{P}$ is a commutative field.*

8. Let $\mathfrak{R}_1$ be a c. o. d. o. i. contained in $\mathfrak{F}$ and $\mathfrak{P}_1$ its prime ideal defined as in the last section. (It is easily seen that $\mathfrak{P}_1$ is the only prime ideal in $\mathfrak{R}_1$). Consider $\mathfrak{R}_2 = \mathfrak{R}(\mathfrak{P}_1)$ the largest d. o. i. in which $\mathfrak{P}_1$ is contained as a left-ideal, i. e., $\mathfrak{R}_2$ consists of the elements a of $\mathfrak{F}$ such that $a\mathfrak{P}_1 \subset \mathfrak{P}_1$. Evidently $\mathfrak{R}_2 \supset \mathfrak{R}_1$ and is c. o. Determine $\mathfrak{R}_3 = \mathfrak{R}(\mathfrak{P}_2)$ where $\mathfrak{P}_2$ is the prime ideal of $\mathfrak{R}_2$ and so on. We thus obtain a sequence of c. o. d. o. i. $\mathfrak{R}_1 \subset \mathfrak{R}_2 \subset \mathfrak{R}_3 \subset \cdots$. Then $\mathfrak{R} = \lim \mathfrak{R}_\nu = \mathfrak{R}_1 \cup \mathfrak{R}_2 \cup \mathfrak{R}_3 \cup \cdots$ is an open d. o. i.

We wish to show that $\mathfrak{R}$ *is compact.* This will follow from the following definition and lemma.

Definition. An element $a \in \mathfrak{F}$ is *integral* if its power sequence $\{a^\nu\}$ has no divergent subsequence.

---

[19] For a proof for hypercomplex systems with a finite basis see L. E. Dickson, *Algebras and their arithmetics,* Chicago, p. 47, or German edition, p. 97.

[20] $\mathfrak{P}^\nu$ is the smallest ring containing all the products $b_1 b_2 \cdots b_\nu$ where $b_\mu \in \mathfrak{P}$.

[21] J. H. M. Wedderburn, "A theorem on finite algebras," *Transactions of the American Mathematical Society,* vol. 6 (1905), pp. 349-352.

LEMMA 8.1. *A subring $\mathfrak{R}$ of $\mathfrak{F}$ containing only integral elements is compact.*

If $\{a_\nu\} \subset \mathfrak{R}$ and $a_\nu \to \infty$ then $a_\nu^{-1} \to 0$. Hence for a sufficiently large N, $a_N^{-1} \in \mathfrak{P}_1$ and so $\lim_{\mu\to\infty} a_N^{-\mu} = 0$. But this implies that $\lim_{\mu\to\infty} a_N^{\mu} = \infty$ which is impossible.

Since $\mathfrak{R} = \lim \mathfrak{R}_\nu$ as well as the $\mathfrak{R}_\nu$ contains integral elements only, it is compact. Since $\mathfrak{R}$ is c. o., it follows by the argument of Theorem 7.1 that $\mathfrak{R} = \mathfrak{R}_N$ for N sufficiently large and hence $\mathfrak{R}(\mathfrak{P}) = \mathfrak{R}$ where $\mathfrak{P}$ is the prime ideal of $\mathfrak{R}$.

Let $\mathfrak{P}^*$ denote the set of inverses of the elements of $\mathfrak{P}$.

THEOREM 8.1. $\mathfrak{F} = \mathfrak{R} \cup \mathfrak{P}^*$.

If $b \in \mathfrak{F}|\mathfrak{R} = \mathfrak{F}|\mathfrak{R}(\mathfrak{P})$, then there exists an element c_1 of $\mathfrak{P}$ such that $b_1 = bc_1 \in \mathfrak{F}|\mathfrak{P}$. If $b_1 \in \mathfrak{F}|\mathfrak{R}$, we may repeat this process and obtain a c_2 in $\mathfrak{P}$ such that $b_1c_2 = bc_1c_2 \in \mathfrak{F}|\mathfrak{P}$ and so on. Since $\mathfrak{P}^\nu \to 0$ the sequence $c_1, c_1c_2, c_1c_2c_3, \cdots$, if infinite, converges to 0 contradicting $bc_1c_2 \cdots c_\nu \in \mathfrak{F}|\mathfrak{P}$. Thus the sequence of c's breaks off after, say, r steps and we obtain $bc_1c_2 \cdots c_r = u \in \mathfrak{R}|\mathfrak{P}$. Hence $b^{-1} = c_1c_2 \cdots c_r u^{-1} \in \mathfrak{P}$, since $u^{-1} \in \mathfrak{R}|\mathfrak{P}$. Thus $\mathfrak{F}|\mathfrak{R} = \mathfrak{P}^*$.

This theorem shows that $\mathfrak{R}$ may be characterized topologically as the totality of integral elements of $\mathfrak{F}$. We have then

THEOREM 8.2. *$\mathfrak{R}$ is the only maximal c. o. d. o. i. in $\mathfrak{F}$.*

If $f \in \mathfrak{P}^\nu|\mathfrak{P}^{\nu+1}$, we say that f has *exponent* ν and if $f^{-1} \in \mathfrak{P}^\nu|\mathfrak{P}^{\nu+1}$, f has *exponent* $-\nu$. Define the value $|f| = \gamma^{\exp f}$ where $f \neq 0$ and $0 < \gamma < 1$ and $|0| = 0$.

THEOREM 8.3. *$|f|$ gives a non-archimedean valuation (Bewertung) of $\mathfrak{F}$: $|f+g| \leqq \max(|f|, |g|)$, $|fg| = |f| \cdot |g|$.*

We choose an element x in $\mathfrak{P}|\mathfrak{P}^2$ and we shall show that if f has exponent ν then $f = f_0x^\nu$ where $f_0 \in \mathfrak{R}|\mathfrak{P}$. First if f and g have exponent ν, fg^{-1} and $g^{-1}f$ have exponent 0. It is sufficient to prove this for $\nu > 0$. Now $fg^{-1} \in \mathfrak{P}$ implies $f \in \mathfrak{P}g \subset \mathfrak{P}^{\nu+1}$ which is contrary to $\exp f = \nu$. Similarly $fg^{-1} \in \mathfrak{P}^*$ is impossible and hence $fg^{-1} \in \mathfrak{R}|\mathfrak{P}$ and $\exp fg^{-1} = 0$. For the same reasons $\exp g^{-1}f = 0$. We observe next that $\exp x^\nu = \nu$. For if $\nu > 0$ $\exp x^\nu \geqq \nu$ and if $> \nu$ then $x^\nu = \Sigma y_1y_2 \cdots y_\mu$ where $y \in \mathfrak{P}$ and $\mu > \nu$. Then $1 = \Sigma x^{-\nu}y_1y_2 \cdots y_\mu$ and since $x^{-1}y \in \mathfrak{R}$ this implies that $1 \in \mathfrak{P}$ which is impossible. Thus $\exp(fx^{-\nu} = f_0) = 0$ and $f = f_0x^\nu$, $f_0 \in \mathfrak{R}|\mathfrak{P}$.

Suppose $f = f_0 x^\nu$, $g = g_0 x^\mu$ where $f_0, g_0 \in \mathfrak{R}|\mathfrak{P}$ and $\nu \leqq \mu$. Then

$$f + g = (f_0 + g_0 x^{\mu-\nu}) x^\nu \qquad (f_0 + g_0 x^{\mu-\nu}) \in \mathfrak{R}.$$

Hence

$$\exp(f + g) \geqq \min(\exp f, \exp g) \qquad |f + g| \leqq \max(|f|, |g|).$$

$$fg = (f_0 x^\nu g_0 x^{-\nu}) x^{\nu+\mu} \qquad (f_0 x^\nu g_0 x^{-\nu}) \in \mathfrak{R}|\mathfrak{P}$$

and so

$$\exp(fg) = \exp f + \exp g \qquad |fg| = |f| \cdot |g|.$$

$\mathfrak{J}$ is a *fractional left-(right-) ideal in* $\mathfrak{F}$ if it is a subgroup and $\mathfrak{R}\mathfrak{J} \subset \mathfrak{J}$ ($\mathfrak{J}\mathfrak{R} \subset \mathfrak{J}$). It follows easily that $\mathfrak{J}$ is open. The inverse $\mathfrak{J}^{-1}$ of $\mathfrak{J}$ is the set of elements a such that $a\mathfrak{J} \subseteq \mathfrak{R}$, or $\mathfrak{J}^{-1}\mathfrak{J} \subset \mathfrak{R}$. $\mathfrak{J}^{-1}$ is evidently a fractional left-ideal. If $\mathfrak{J} \neq (0)$, $\neq 1$ then $\mathfrak{J}$ is called *proper*.

THEOREM 8.4. *Every proper fractional left-(right-) ideal is two-sided, principal and a power of* $\mathfrak{P}$.

If $b \in \mathfrak{J}$ and c has exponent $\geqq \exp b$, then $\exp cb^{-1} \geqq 0$, and $cb^{-1} \in \mathfrak{R}$. Hence $c \in \mathfrak{R}b \subset \mathfrak{J}$. If $\mathfrak{J}$ contains elements of arbitrarily small exponent, $\mathfrak{J} = \mathfrak{F}$. Otherwise let b have the smallest exponent of the elements of $\mathfrak{J}$. It follows that $\mathfrak{J}$ is the set of elements of $\mathfrak{F}$ of exponent $\geqq \exp b = k$ and hence $\mathfrak{J} = \mathfrak{P}^k = (b)$.

It follows directly from this theorem that $\mathfrak{J}\mathfrak{J}^{-1} = \mathfrak{R}$ as well as $\mathfrak{J}^{-1}\mathfrak{J}$. If $\mathfrak{F}$ has characteristic 0, $p\mathfrak{R}$ is an ideal $\neq 0$ and hence is $= \mathfrak{P}^e$ where $e \geqq 1$.

9. In this final section we shall apply the arithmetic results obtained in § 8 to obtain the structure of $\mathfrak{F}$.

Let p^n be the order of the finite field $\mathfrak{R} - \mathfrak{P}$ of characteristic p. $\mathfrak{R} - \mathfrak{P}$ contains a primitive q-th root of unity where $q = p^n - 1$, i. e., there exists a $u_0 \in \mathfrak{R}$ such that $u_0^q \equiv 1(\mathfrak{P})$ and q is the smallest power for which such a congruence holds. Let $x \equiv 0(\mathfrak{P})$ but $\not\equiv 0(\mathfrak{P}^2)$. Then the correspondence $a \leftrightarrow xax^{-1}$ where $a \in \mathfrak{R}$ determines an automorphism of the Galois group of $\mathfrak{R} - \mathfrak{P}$. It follows that

$$xu_0x^{-1} \equiv u_0^s(\mathfrak{P}), \quad \text{where} \quad s = p^t.$$

THEOREM 9.1. $\mathfrak{F}$ *contains a primitive q-th root of unity.*

We shall determine $u \in \mathfrak{R}|\mathfrak{P}$ such that $u^q = 1$ and $u \equiv u_0(\mathfrak{P})$. Since u_0 is a primitive q-th root of unity mod $\mathfrak{P}$ it will follow that u is primitive. Let r be the exponent of s mod q, i. e., r is the least positive integer for which $s^r \equiv 1(q)$. Then r is also the least positive integer satisfying

$$(**) \qquad x^r u_0 x^{-r} \equiv u_0(\mathfrak{P}).$$

If $u_0{}^q \neq 1$, suppose $u_0{}^q \equiv 1(\mathfrak{P}^k)$, $\not\equiv 1(\mathfrak{P}^{k+1})$. Then $u_0{}^q - 1 = v_0 x^k$ where $v_0 \in \mathfrak{R}|\mathfrak{P}$. Since $v_0 u_0 \equiv u_0 v_0(\mathfrak{P})$ and $(u_0{}^q - 1)u_0 = u_0(u_0{}^q - 1)$, we have $x^{-k}u_0 x^k \equiv u_0(\mathfrak{P})$ and hence by $(**)$ $k \equiv 0(r)$, say $k = rl$. Thus

$$u_0{}^q \equiv 1 + wx^{rl}(\mathfrak{P}^{rl+1}) \qquad w \not\equiv 0(\mathfrak{P}).$$

Set $u_1 = u_0 + yx^{rl}$ where y is to be determined $\in \mathfrak{R}$ so that $u_1{}^q \equiv 1(\mathfrak{P}^{rl_1})$ for $l_1 > l$. This requires in particular that

$$({}_*{}^*{}_*) \qquad u_1{}^q \equiv 1 + (w + qu_0{}^{q-1}y)x^{rl} \equiv 1(\mathfrak{P}^{rl+1})$$

by $(**)$ and the commutativity of u_0 and y mod $\mathfrak{P}$. Since q and $u_0{}^{q-1} \not\equiv 0(\mathfrak{P})$, y may be chosen so that $(w + qu_0{}^{q-1}y) \equiv 0(\mathfrak{P})$ and will then satisfy $({}_*{}^*{}_*)$. But then we will necessarily have $u_1{}^q \equiv 1(\mathfrak{P}^{rl_1})$ for $l_1 > l$. If u_1 is not a q-th root of one we may repeat this process with it in place of u_0. In this way we either obtain a q-th root of one after a finite number of steps or else an infinite sequence $\{u_\nu\}$ such that

$$u_\nu \equiv u_{\nu-1}(\mathfrak{P}^{rl_{\nu-1}}) \qquad u_\nu{}^q \equiv 1\,(\mathfrak{P}^{rl_\nu})$$

where $l < l_1 < l_2 < \cdots$. The $\{u_\nu\}$ converge to a limit u having the desired properties:

$$u^q = 1 \qquad u \equiv u_\nu(\mathfrak{P}^{rl_{\nu-1}}) \qquad xux^{-1} \equiv u^s(\mathfrak{P}).$$

THEOREM 9.2. *x may be normalized so that $xux^{-1} = u^s$, $x \in \mathfrak{P}|\mathfrak{P}^2$.*

This theorem has been given by Hasse for p-adic fields. Moreover his proof (H. p. 511) goes over word for word for the present more general case.

THEOREM 9.3. *Every element of $\mathfrak{F}$ may be represented uniquely in the form $v = (v_0 + v_1 x + v_2 x^2 + \cdots)x^{-k}$ where $k \geqq 0$ and $v_\nu = u^{k_\nu}$ or 0.*

If $v \equiv 0(\mathfrak{P}^{-k})$ for $k \geqq 0$, then $vx^k \equiv 0(\mathfrak{R})$ and hence $vx^k \equiv v_0(\mathfrak{P})$ where $v_0 = u^{k_0}$ or 0. Now $(vx^k - v_0)x^{-1} \equiv v_1(\mathfrak{P})$ where $v_1 = u^{k_1}$ or 0 and so on. We obtain thus a sequence $v_\nu = u^{k_\nu}$ or 0 such that

$$vx^k \equiv v_0 + v_1 x + v_2 x^2 + \cdots + v_\nu x^\nu (\mathfrak{P}^{\nu+1})$$

and hence

$$v = (v_0 + v_1 x + v_2 x^2 + \cdots)x^{-k}.$$

If $(\sum_{\nu=0}^{\infty} v_\nu x^\nu)x^{-k} = (\sum_{\nu=0}^{\infty} v'_\nu x^\nu)x^{-l}$, evidently $k = l$ so that $\Sigma v_\nu x^\nu = \Sigma v'_\nu x^\nu$. Hence

$v_0 \equiv v'_0(\mathfrak{P})$ and so $v_0 = v'_0$ since v_0 and v'_0 are members of the complete set of incongruent residues $0, u, u^2, \cdots, u^q \bmod \mathfrak{P}$. Likewise $v_1 = v'_1, v_2 = v'_2, \cdots$.

THEOREM 9.4. *The centrum $\mathfrak{C}$ of $\mathfrak{F}$ consists of all elements of the form $(c_0 + c_1x^r + c_2x^{2r} + \cdots)x^{-kr}$ where $c_\nu = 0$ or $u^{k\nu}$ and commutes with x.*

If $c = \sum_{\nu=-k}^{\infty} c_\nu x^\nu$, $u^{-1}cu = c$ implies $c_\nu = c_\nu u^{s^\nu - 1}$. Hence if $c_\nu \neq 0$, $u^{s^\nu-1} = 1$ and $s^\nu \equiv 1(q)$ or $\nu \equiv 0(r)$. Further $xcx^{-1} = c$ implies that $c_\nu{}^s = c_\nu$ and hence c_ν commutes with x. Thus the two conditions that c commutes with x and with u imply that c has the form $\sum_{\nu=-k}^{\infty} c_\nu x^{\nu r}$, $c_\nu x = xc_\nu$. Since these two conditions insure that c belongs to the centrum $\mathfrak{C}$, we have the theorem.

The above proof shows also that the elements of $\mathfrak{F}$ commutative with u are all of the form $\sum_{\nu=-k}^{\infty} v_\nu x^{r\nu}$ where $v_\nu = u^{k\nu}$ or 0. These elements form a commutative subfield $\mathfrak{C}' \supset \mathfrak{C}$. Since $\mathfrak{C}' = \mathfrak{C}(u)$ it has a finite basis over $\mathfrak{C}$.

THEOREM 9.5. *$\mathfrak{C}'$ is a cyclic field of degree r over $\mathfrak{C}$.*

The automorphism $a \leftrightarrow xax^{-1}$ of $\mathfrak{F}$ leaves $\mathfrak{C}'$ invariant (though not element-wise) and hence induces an automorphism S in $\mathfrak{C}'$. The elements of $\mathfrak{C}'$ invariant under S are precisely the elements commutative with x and hence belonging to $\mathfrak{C}$. Since x^r is the smallest power of x commutative with u, the order of S is r. Hence $\mathfrak{C}'$ is cyclic of degree r over $\mathfrak{C}$ with S as generating automorphism of its Galois group relative to $\mathfrak{C}$.

THEOREM 9.6. *$\mathfrak{F}$ is a cyclic algebra over its centrum.*

For, by Theorem 9.5, u satisfies an irreducible equation of degree r, $\phi_r(u) = 0$ having coefficients in $\mathfrak{C}$. Thus

$$\phi_r(u) = 0 \qquad xu = u^s x \qquad (s = p^t) \qquad x^r = z \in \mathfrak{C}$$

gives a description of the algebra $\mathfrak{F}$ relative to its centrum $\mathfrak{C}$.

To complete the description of $\mathfrak{F}$ it is necessary to give the structure of $\mathfrak{C}$. Since $\mathfrak{C}$ is closed it is a l. c. s. t. d. commutative field and hence as mentioned in the introduction its form has been described by v. Dantzig and by Hasse and Schmidt. We shall merely state the results here and sketch briefly their derivation.

For the case $\chi(\mathfrak{F}) = 0$ we have seen that $\mathfrak{F}$ and hence also $\mathfrak{C}$ has finite order over the p-adic field P_p. From considerations analogous to those of

Theorems 9.1 and 9.2 it follows that $\mathfrak{C}$ is generated by a $p^m - 1$ ($m \geqq 1$) root of one and a second element z ($= x^r$). Certain additional normalizations may be made. The reader is referred to Hasse's paper (H. p. 514) for these.

If $\chi(\mathfrak{F}) = p$, we have

THEOREM 9.7. *The powers of u and 0 form a finite field K_{p^n} of p^n elements.*

It suffices to show that $v \equiv u^j + u^k = u^l$ or 0 and $-u^k = u^m$. Suppose $v \neq 0$. Since v is algebraic mod p (satisfies an algebraic equation whose coefficients are in the field K_p of residues mod p), it is a root of unity and hence $\epsilon\, \mathfrak{R}|\mathfrak{P}$. If $v \equiv u^l \pmod{\mathfrak{P}}$, $v - u^l$ is algebraic mod p and $\equiv 0 \pmod{\mathfrak{P}}$. This is impossible unless $v - u^l = 0$. Similarly we may show that $-u^k$ is $=$ some u^m.

By the same argument we see that x is transcendental mod p. Thus $\mathfrak{F}$ is uniquely determined by the field K_{p^n} and its automorphism $a \leftrightarrow a^s$ ($s = p^t$). For when these are given we have $\mathfrak{F}$ determined as the set of integral power series in x with a finite number of negative powers and coefficients in K_{p^n} where the multiplication is determined by $xa = a^s x (a \epsilon K_{p^n})$.

PRINCETON UNIVERSITY.

ON SINGULAR FOURIER-STIELTJES TRANSFORMS.

By Richard Kershner.

Let $\sigma(x)$, where $-\infty < x < +\infty$, be continuous, bounded and monotone non-decreasing and not a constant. If $\mu(S)$ denotes the measure of the set S of points x at which $\sigma(x)$ is actually increasing then $\mu(S) > 0$ is a necessary condition for the absolute continuity of $\sigma(x)$. If $\sigma(x)$ is absolutely continuous, then

$$(1) \qquad L(t,\sigma) \to 0 \text{ as } t \to \pm\infty, \text{ where } L(t,\sigma) = \int_{-\infty}^{\infty} e^{itx} d\sigma(x).$$

The question, under what additional condition does (1) imply the absolute continuity of $\sigma(x)$, is related to the problem of sets of uniqueness in the theory of trigonometrical series;[1] hence it is to be expected that the absolute continuity of σ must depend on arithmetical details. The present note furnishes among other things, an illustration to this situation.

It is known[2] that there exists for every positive $a < 1$ a continuous monotone $\sigma(x) = \sigma_a(x)$ such that

$$(2) \qquad L(t,\sigma_a) = \prod_{n=1}^{\infty} \cos(a^n t), \qquad -\infty < t < +\infty,$$

and that[3] the measure $\mu(S)$ of the corresponding set $S = S_a$ is zero whenever $0 < a < 1/2$. Now it will be shown that if a is a rational number in the latter interval, then (1) is satisfied by $\sigma = \sigma_a$ if and only if a^{-1} is not an integer. More precisely

$$(3) \quad L(t,\sigma_a) = O[(\log|t|)^{-\gamma}], \quad \gamma = \frac{-\log\cos(\pi/2q)}{\log\,(2\log q/\log p)} > 0,\ t \to \pm\infty$$

if $0 < a = p/q < 1/2$, where p and q are relatively prime and $p > 1$. In particular *there exist continuous monotone bounded functions σ which satisfy* (1) *and $\mu(S) = 0$ and are not absolutely continuous.* This implies a result of Menchoff[4] who considered the corresponding question for the case of the Fourier-Stieltjes coefficients where $t \to \pm\infty$ through integral values, rather than continuously.

---

[1] Cf. A. Zygmund, *Trigonometrical Series* (1935), p. 291 et seq.

[2] Cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[3] Cf. R. Kershner and A. Wintner, "On symmetric Bernoulli convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 541-548.

[4] Cf. A. Zygmund, *ibid.*

If $a = 1/3$, then σ_a is the Cantor function for which it is known that (1) does not hold.[5] That (1) does not hold for $\sigma = \sigma_a$ $(a = 1/4, 1/5, \cdots)$, can be shown in exactly the same way as is usually done for the Cantor function.[6] Hence it may be assumed that $p > 1$. In the proof of (3) it will not be necessary to assume that $0 < a < 1/2$ but only $0 < a < 1$, i. e. $q > p$. It is clearly sufficient to consider $t > 0$ since $L(t, \sigma_a)$ is an even function.

Let λ denote the set of all points $t > 0$ which are within a distance $\pi/2q$ of an integral multiple of π so that

$$|\cos t| > \cos \pi/2q$$

if and only if t is in λ. Let $t > \pi$ and let $h_0 = h_0(t) \geqq 0$ be the unique integer for which

$$q^{h_0+1} > t/\pi \geqq q^{h_0}. \tag{4}$$

Now either t is not in λ or there is a unique integer $k = k(t)$ such that

$$t = k\pi + \delta, \qquad |\delta| < \pi/2q \tag{5}$$

where $k < q^{h_0+2}$ by definition of h_0. In the latter case write

$$k = q^J l \tag{6}$$

where $J = J(t)$ is the exponent of the highest power of q contained in k, so that $0 \leqq J \leqq h_0 + 1$ and q does not divide l. By (5) and (6) we have

$$t(p/q)^{J+1} = \pi p^{J+1} l/q + \delta(p/q)^{J+1}.$$

Since p and q are relatively prime and q does not divide l then $p^{J+1}l/q$ must differ from an integer by at least $1/q$. And since

$$|\delta(p/q)^{J+1}| < |\delta| < \pi/2q$$

we see that $t(p/q)^{J+1}$ differs from an integral multiple of π by at least $\pi/q - \pi/2q = \pi/2q$ and hence that $t(p/q)^{J+1}$ is not in λ. Thus for any t satisfying (4) we know that $t(p/q)^{m_0}$ is not in λ for some m_0 such that $0 \leqq m_0 \leqq h_0 + 1$. Furthermore from (4) we get

$$p^{h_0+2}/q > t(p/q)^{h_0+2}/\pi \geqq p^{h_0+2}/q^2. \tag{7}$$

Now let t be chosen so large, for a fixed $n > 2$, that $h_0 = h_0(t)$ satisfies

$$h_0 + 2 \geqq (2 \log q/\log p)^n$$

[5] Cf. B. Jessen and A. Wintner, *ibid.*, where further references are given.

[6] Note that this proof breaks down in the case $a = 1/2$ since $\cos \pi/2 = 0$ and in fact in this case (1) holds. Cf. B. Jessen and A. Wintner, *ibid.*

this choice of t being possible in view of (4). Then

$$(h_0+2)(\log p/\log q) \geqq 2(2\log q/\log p)^{n-1} > (2\log q/\log p)^{n-1}+1.$$

Hence we can find an integer $g = g(t) \geqq 0$ such that

$$(h_0+2)(\log p/\log q) \geqq g+2 \geqq (2\log q/\log p)^{n-1}.$$

The first of these inequalities may be written

$$p^{h_0+2}/q^2 \geqq q^g$$

and comparing this with (7) we see that there is a unique integer $h_1 = h_1(t)$ such that

$$q^{h_1+1} > t(p/q)^{h_0+2}/\pi \geqq q^{h_1}$$

and

$$(h_1+2) \geqq (g+2) \geqq (2\log q/\log p)^{n-1}.$$

Applying to $t(p/q)^{h_0+2}$ the argument that we applied above to t we see that $t(p/q)^{m_1}$ is not in λ for some m_1 such that $h_0+2 \leqq m_1 \leqq h_0+h_1+3$ and that

$$q^{h_2+1} \geqq t(p/q)^{h_0+h_1+4}/\pi \geqq q^{h_2}$$

where

$$h_2+2 \geqq (2\log q/\log p)^{n-2}.$$

Applying the above argument $n-1$ times it follows that if

(8) $$t/\pi > q^h \quad \text{where} \quad h > (2\log q/\log p)^n$$

then there are at least $n-1$ distinct values of m such that

$$0 \leqq m \leqq h_0+h_1+\cdots+h_{n-1}+2^{n-1}-1$$

and that $t(p/q)^m$ is not in λ and consequently

(9) $$|L(t,\sigma_a)| < \cos^{n-1}(\pi/2q).$$

Let

(10) $$t/\pi \geqq q^{A^{n+1}} \quad \text{where} \quad A = A_{pq} = (2\log q/\log p) > 2.$$

Then certainly there is an h satisfying (8).

Let $n(t)$ be the largest integer satisfying (10). Then

$$n(t) > \{\log[\log t - \log\pi] - \log\log q - 2\log A\}/\log A$$

and substitution into (9) gives the required appraisal (3).

THE JOHNS HOPKINS UNIVERSITY.

THE RELATION OF THE CLASSICAL ORTHOGONAL POLYNOMIALS TO THE POLYNOMIALS OF APPELL.

By J. Shohat.

Introduction. The classical orthogonal polynomials ($\equiv COP$) of Jacobi (J), Laguerre (L) and Hermite (H) all satisfy a differential equation of the following type: [1]

$$A(x)y'' + B(x)y' + C_n y = 0,$$

$$(J) \text{ in } (0,1): \quad x^{\alpha-1}(1-x)^{\beta-1}T_n'' + [\alpha-(\alpha+\beta)x]T_n' + n(n+\alpha+\beta-1)T_n = 0,$$

$$T_n \equiv T_n(x;\alpha,\beta), \quad \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}T_mT_n dx = 0;$$

$$xL_n'' + (\alpha - x)L_n' + nL_n = 0,$$

$$(L) \qquad L_n \equiv L_n(x;\alpha), \quad \int_0^\infty x^{\alpha-1}e^{-x}L_mL_n dx = 0;$$

$$(H) \qquad H_n'' - 2xH_n' + 2nH_n = 0, \quad H_n \equiv H_n(x),$$

$$\int_{-\infty}^\infty e^{-x^2}H_mH_n dx = 0;$$

$$(\alpha, \beta > 0;\ m, n = 0, 1, \cdots;\ m \neq n).$$

They all enjoy the property that their derivatives again form orthogonal systems of polynomials, with the new weight-function

$$(2) \quad p_1(x) = A(x)p(x);\ p(x) = x^{\alpha-1}(1-x)^{\beta-1}(J),\ x^{\alpha-1}e^{-x}(L),\ e^{-x^2}(H).$$

Thus, taking the coefficient of x^n equal to unity and using for orthogonal polynomials ($\equiv OP$) with the weight-function $p(x)$ the notation

$$(3) \qquad \Phi_n(x;p) \equiv \Phi_n(x) = x^n - S_n x^{n-1} + d_{n,n-2}x^{n-2} + \cdots \qquad (n = 0, 1, \cdots),$$

we have for COP

$$(4) \qquad \Phi'_n(x;p) = n\Phi_{n-1}(x;Ap) \qquad (n = 0, 1, \cdots).$$

For Hermite polynomials (4) takes the simplest form

$$(5) \qquad \Phi'_n(x) = n\Phi_{n-1}(x) \qquad (n = 0, 1, \cdots).$$

---

[1] The notations here employed are the same as those in my work: "Théorie générale des polynomes orthogonaux de Tchebichef," *Mémorial des Sciences Mathématiques*, Fasc. 66. The reader is referred to this work for further details concerning theorems and formulae.

(4, 5) naturally lead to the study of the relation between *COP* and Appell[2] polynomials ($\equiv AP$) for which the general definition is

$$(6) \qquad A'_n(x) = nA_{n-1}(x)$$

($A_n(x)$, a polynomial of degree $n = 0, 1, \cdots$; thus $A_n(x) = x^n + \cdots$).

This study is the main object of the present paper. It is based upon the difference equation *characteristic for OP* (see Theorem I below):

$$(7) \qquad \Phi_n(x) = (x - c_n)\Phi_{n-1}(x) - \lambda_n\Phi_{n-2}(x)$$
$$(n \geqq 2;\ \Phi_0 = 1,\ \Phi_1(x) = x - c_1;\ c_n,\ \lambda_n\text{-const.}).$$

In this connection we give some interesting properties of the constants c_n, λ_n in (7), also some limitations of the zeros of $\Phi_n(x)$. We conclude with a new simple proof that Hermite polynomials form the only sequence of polynomials which are at the same time *AP* and *OP*.

I. *The difference equation for OP.*

THEOREM I. *A necessary and sufficient condition that the sequence of polynomials*

$$\Phi_0(x) = 1,\ \ \Phi_1(x) = x - c_1,\ \ \Phi_n(x) = x^n - S_n x^{n-1} + d_{n,n-2}x^{n-2} + \cdots$$
$$(n = 2, 3, \cdots)$$

form a sequence of OP is that they satisfy a difference equation of the form

$$(7) \qquad \Phi_n(x) = (x - c_n)\Phi_{n-1}(x) - \lambda_n\Phi_{n-2}(x)$$
$$(n \geqq 2;\ c_n, \lambda_n \text{ are constants})$$

with positive λ_n.

Proof. The necessity of (7) is well known. Its sufficiency, in view of the importance of the theorem, will be proved in extenso as follows.[3] (7) implies that $\{\Phi_n\}$ are denominators of the successive convergents $\Omega_n(x)/\Phi_n(x)$ $(n = 0, 1, \cdots)$ of the continued fraction

$$(8) \qquad F(x) \equiv \frac{\lambda_1/}{/x - c_1} - \frac{\lambda_2/}{/x - c_2} - \cdots - \frac{\lambda_n/}{/x - c_n} - \cdots$$
$$(\lambda_1 > 0, \text{ arbitrary}),$$

whence the characteristic property:

---

[2] Appell, "Sur une classe de polynomes," *Ann. Ec. Norm.*, 2, vol. 9 (1880).

[3] We have been in possession of this proof for several years. Recently J. Favard published an identical proof in the *Comptes Rendus* ("Sur les polynomes de Tchebicheff," *Comptes Rendus*, vol. 200 (1935)). Cf. also, O. Perron, *Die Lehre von den Kettenbrüchen*, 2 ed., 377 ff., and J. Sherman, "On the numerators of the convergents of the Stieltjes continued fractions," *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 64-87.

(9) $$F(x)-\frac{\Omega_n(x)}{\Phi_n(x)}=\frac{\alpha'}{x^{2n+1}}+\cdots\equiv\left(\frac{1}{x^{2n+1}}\right) \qquad (\alpha',\cdots \text{ constants}).$$

It follows that

(9 bis) $$\frac{\Omega_n(x)}{\Phi_n(x)}=\frac{\alpha_0}{x}+\frac{\alpha_1}{x^2}+\cdots+\frac{\alpha_{2n}}{x^{2n}}+\frac{\beta}{x^{2n+1}}+\cdots \qquad (\alpha_0=\lambda_1),$$
$$\frac{\Omega_{n+1}(x)}{\Phi_{n+1}(x)}=\frac{\alpha_0}{x}+\cdots+\frac{\alpha_{2n}}{x^{2n}}+\frac{\alpha_{2n+1}}{x^{2n+1}}+\frac{\alpha_{2n+2}}{x^{2n+2}}+\frac{\gamma}{x^{2n+3}}+\cdots,$$

and the formal expansion of $F(x)$ in ascending powers of $1/x$ is

$$F(x)=\frac{\alpha_0}{x}+\frac{\alpha_1}{x^2}+\cdots+\frac{\alpha_n}{x^n}+\cdots \qquad (\alpha_0=\lambda_1>0).$$

The theory of continued fractions enables us to conclude from the existence of convergents for (8) of *all* ranks $n=0,1,\cdots$, that all determinants

(10) $$\Delta_n=\begin{vmatrix} \alpha_0 & \alpha_1\cdots\alpha_{n-1} \\ \alpha_1 & \alpha_2\cdots\alpha_n \\ \cdot & \cdot\quad\cdot\quad\cdot\quad\cdot \\ \cdot & \cdot\quad\cdot\quad\cdot\quad\cdot \\ \alpha_{n-1} & \alpha_n\cdots\alpha_{2n-2} \end{vmatrix}\neq 0 \quad (n=0,1,\cdots;\ \Delta_0=1;\ \Delta_1=\alpha_0).$$

Moreover,

(11) $$\lambda_n=\frac{\Delta_{n-2}\Delta_n}{\Delta^2_{n-1}} \qquad (n\geqq 2;\ \lambda_1=\alpha_0=\Delta_1).$$

Since, by hypothesis, all $\lambda_n>0$ $(n\geqq 2)$ and λ_1 $(\equiv\Delta_1)$ is also chosen positive, we see at once that *all* Δ_n *are positive* $(n=0,1,\cdots)$. But this is Hamburger's [4] condition for the existence of (at least) one monotonic bounded function $\psi(x)$, non-decreasing in $(-\infty,\infty)$, a solution of the Moments Problem

(12) $$\alpha_n\ (\equiv \text{moment of order } n)=\int_{-\infty}^{\infty}x^n d\psi(x) \qquad (n=0,1,2,\cdots).$$

(9), rewritten as

(13) $$\Phi_n(x)F(x)=\Omega_n(x)+(1/x^{n+1})$$
$$(\Phi_n(x)=x^n+f_1x^{n-1}+\cdots+f_n),$$

leads to the system of equations

$$f_n\alpha_i+f_{n-1}\alpha_{i+1}+\cdots+f_1\alpha_{i+n-1}+\alpha_{i+n}=0 \qquad (i=0,1,\cdots,n-1),$$

which, by virtue of (12), can be written in integral form as

[4] H. Hamburger, "Ueber eine Erweiterung des Stieltjesschen Momentenproblem, I, II, III," *Mathematische Annalen*, vols. 81, 82 (1920-1921), pp. 235-319, 120-164, 168-187.

$$(14)\qquad \int_{-\infty}^{\infty} \Phi_n(x) x^i d\psi(x) = 0 \qquad (0 \leqq i < n = 1, 2, \cdots),$$

which is one of the expressions of the orthogonality property of $\{\Phi_n(x)\}$.

Remarks. (i) The interval $(-\infty, \infty)$ in (14) may be reduced to a subinterval (a, b), if $\psi(x)$ remains constant outside (a, b). The "true" interval of orthogonality is given by the limiting values (finite or infinite, but always existing) of the two extreme zeros of $\Phi_n(x)$, as $n \to \infty$. Hereafter, "interval of orthogonality" means the "true" interval. (ii) The Moments Problem (12) may have more than one solution. All such solutions evidently generate the same sequence $\{\Phi_n(x)\}$ of *OP*, so that from this standpoint they are all equivalent. We shall write

$$(15)\qquad \Phi_n(x; d\psi) \equiv \Phi_n(x) \equiv \Phi_n.$$

2. *Symmetric OP.* By this we mean $\{\Phi_n(x)\}$ with the property

$$(16)\qquad \Phi_n(-x) \equiv (-1)^n \Phi_n(x) \qquad (n = 1, 2, \cdots).$$

(Legendre or Hermite polynomials are illustrations.) We have the following obviously necessary and sufficient condition that the sequence $\{\Phi_n(x)\}$ be symmetric, namely: *all* c_n $(n \geqq 2)$ *in the recurrence relation* (7), *also* c_1, *vanish,* and this is equivalent to the vanishing of S_n— sum of the zeros of $\{\Phi_n(x)\}$ $(n = 1, 2, \cdots)$. In fact,

$$(17)\qquad S_n = c_1 + c_2 + \cdots + c_n \qquad (S_n - S_{n-1} = c_n).$$

The following remark will be used later: if in (7)

$$(18)\qquad c_1 = c_2 = \cdots = c_n = \cdots = c,$$

then $\{\Phi_n(x)\}$ is reducible to a symmetric sequence by the linear substitution $(x - c) | x$. It is of interest to interpret the symmetric property (16) in terms of the moments $\{\alpha_n\}$. This is given in

THEOREM II. *The sequence* $\{\Phi_n\}$ *is symmetric if and only if all odd moments* $\alpha_1, \alpha_3, \cdots$ *vanish. It is generated by* $\psi(x)$ *with the property:* $\psi(-x) \equiv -\psi(x)$ *in* $(-\infty, \infty)$.

Proof. Assuming (16), we conclude that the continued fraction (8), for which Ω_n/Φ_n is the $(n+1)$-st convergent, takes the form

$$\frac{\lambda_1/}{/x} - \frac{\lambda_2/}{/x} - \cdots - \frac{\lambda_n/}{/x} - \cdots,$$

so that $\Omega_n(x)$, of degree $n - 1$ (satisfying the same recurrence relation (7), with $\Omega_0 = 0$, $\Omega_1 = \lambda_1$) is also symmetric. Hence,

$$\frac{\Omega_{2n}(x)}{\Phi_{2n}(x)}=\frac{xP_{n-1}(x^2)}{Q_n(x^2)},\qquad \frac{\Omega_{2n+1}(x)}{\Phi_{2n+1}(x)}=\frac{P_n(x^2)}{xS_n(x^2)}\qquad (n=1,2,\cdots),$$

(P_i, Q_i, S_i are polynomials of degree i),

and this, by virtue of (9, 9 bis, 12), shows that

$$\alpha_{2n-1}=\int_{-\infty}^{\infty}x^{2n-1}d\psi(x)=0\qquad (n=1,2,\cdots). \tag{19}$$

Moreover,

$$\alpha_{2n}=\int_{-\infty}^{\infty}x^{2n}d\psi(x)=-\int_{-\infty}^{\infty}x^{2n}d\psi(-x); \tag{20}$$
$$\alpha_{2n+1}=\int_{-\infty}^{\infty}x^{2n+1}d\psi(x)=\int_{-\infty}^{\infty}x^{2n+1}d\psi(-x).$$

Now consider the function

$$\phi(x)\equiv\frac{\psi(x)-\psi(-x)}{2},$$

which is also monotonic non-decreasing in $(-\infty,\infty)$. From (20),

$$\int_{-\infty}^{\infty}x^{2n}d\phi(x)=\int_{-\infty}^{\infty}x^{2n}d\psi(x),\quad \int_{-\infty}^{\infty}x^{2n+1}d\phi(x)=0\qquad (n=0,1,\cdots).$$

It follows from (19), that $\phi(x)$ has the same moments as $\psi(x)$. Thus $\{\Phi_n(x)\}$ may be considered as $\{\Phi_n(x;d\phi)\}$, where evidently

$$\phi(-x)\equiv-\phi(x)\ \text{ in }\ (-\infty,\infty), \tag{21}$$

and the necessity of the condition stated is established. That it is also sufficient, we prove as follows: If all $\alpha_{2n+1}=0$ $(n=0,1,\cdots)$, we conclude as above that $\Phi_n(x)\equiv\Phi_n(x;d\psi)$ with $\psi(-x)\equiv-\psi(x)$. The orthogonality property

$$\int_{-\infty}^{\infty}\Phi_n(x)G_{n-1}(x)\,d\psi(x)=0\qquad (n=1,2,\cdots),$$

where $G_s(x)$ generally stands for an *arbitrary* polynomial of degree $\leqq s$, gives now

$$\int_{-\infty}^{\infty}\Phi_n(-x)G_{n-1}(-x)\,d\psi(x)=0,$$

and this shows (in view of the uniqueness, to within constant factors, of the sequence $\{\Phi_n(x)\}$) that

$$\Phi_n(-x)\equiv\text{Const.}\cdot\Phi_n(x)\equiv(-1)^n\Phi_n(x).^{5}$$

[5] Another proof of Theorem II, based on the determinant-representation of $\Phi_n(x)$ was communicated to me by Mr. G. Milgram.

3. *Some properties of the constants c_n, λ_n.* It is interesting to note that, given an arbitrary sequence of real constants $\{c_n, \lambda_n\}$ $(n = 1, 2, \cdots)$ subject to the single limitation $\lambda_n > 0$ for all n, we get, by means of (7), a sequence of OP. Thus we get a new approach to the theory of OP through the study of the sequences $\{c_n\}, \{\lambda_n\}$. This is clearly shown in the discussion which follows.

THEOREM III. (i) *The interval of orthogonality for $\{\Phi_n(x)\}$ is finite if and only if both sequences $\{c_n\}, \{\lambda_n\}$ are bounded; in other words, if any one of these sequences is unbounded, the interval of orthogonality is infinite.* (ii) *Both sequences are unbounded, if the orthogonality interval is* $(0, \infty)$. (iii) *If only one of the sequences $\{c_n\}, \{\lambda_n\}$ is bounded, the orthogonality interval is then* $(-\infty, \infty)$. [(a, ∞) or $(-\infty, a)$, a finite, can be reduced to $(0, \infty)$ by a linear transformation].

Proof. (i) Necessity follows from the inequalities

$$a < c_n < b;\ \lambda_n < (b - a)^2/4,$$

and sufficiency—from the fact that the zeros of the sequence $\{\Phi_n(x)\}$ all lie within finite limits. We prove the latter statement by making very simple use of the recurrence relation (7) as follows. We have

$$\Phi_1(x) = x - c_1 > 0, \text{ for } x > c_1,$$

$$\frac{\Phi_2(x)}{\Phi_1(x)} = x - c_2 - \frac{\lambda_2}{x - c_1} \geqq k_2 > 0,$$

$$\text{for } x \geqq X_2 = \frac{c_1 + c_2 + k_2 + \sqrt{(c_1 + c_2 + k_2)^2 - 4(c_1k_2 + c_1c_2 - \lambda_2)}}{2}$$

$$\text{and } x > c_1,$$

$$\frac{\Phi_3(x)}{\Phi_2(x)} = x - c_3 - \frac{\lambda_3}{\phi_2/\phi_1} > x - c_3 - \frac{\lambda_3}{k_2} \geqq k_3 > 0,$$

$$\text{for } x \geqq c_3 + \frac{\lambda_3}{k_2} + k_3,$$

$$\text{and } x > c_1,\ x \geqq X_2,$$

. . . .

$$\frac{\Phi_n(x)}{\Phi_{n-1}(x)} \geqq k_n > 0, \text{ for } x > \max\left(c_1, X_2, c_i + \frac{\lambda_i}{k_{i-1}} + k_i\right), \quad (i = 2, 3, \cdots, n).$$

Similarly we treat the values of x for which

$$\Phi_1(x) < 0, \frac{\Phi_2(x)}{\Phi_1(x)} < 0, \cdots, \frac{\Phi_n(x)}{\Phi_{n-1}(x)} < 0.$$

We thus conclude that *the zeros $x_{1,n} < x_{2,n} < \cdots < x_{n,n}$ of $\Phi_n(x)$ lie in the following interval:*

$$\left(\min\left[c_1, X_1, c_i + k'_i + \frac{\lambda_i}{k'_{i-1}}\right], \max\left[c_1, X_2, c_i + k_i + \frac{\lambda_i}{k_{i-1}}\right]\right)$$

(22) $\qquad k_i > 0,\ k'_i < 0$, arbitrary; $i = 2, 3, \cdots n$

$$X_1 = \frac{c_1 + c_2 + k'_2 - \sqrt{(c_1 + c_2 + k_2)^2 - 4(c_1 k'_2 + c_1 c_2 - \lambda_2)}}{2} \equiv X_1(c_1, c_2, k'_2);$$

$$X_2 = \frac{c_1 + c_2 - k_2 + \sqrt{\cdots}}{2} \equiv X_2(c_1, c_2, k_2).$$

Taking

$$-k'_i = k_i = \sqrt{\lambda_{i+1}}\,(i = 2, 3, \cdots, n-1), -k'_n = k_n = \sqrt{\lambda_n},$$

we conclude that

$$\begin{aligned}(23)\qquad &\min(c_1, X_1, c_i - \sqrt{\lambda_i} - \sqrt{\lambda_{i+1}}, c_n - 2\sqrt{\lambda_n}) < x_{1,n} < x_{n,n}\\ &< \max(c_1, X_2, c_i + \sqrt{\lambda_i} + \sqrt{\lambda_{i+1}}, c_n + 2\sqrt{\lambda_n})\\ &\qquad (i = 2, 3, \cdots, n-1)\\ &X_1 = X_1(c_1, c_2, -\sqrt{\lambda_3}), X_2 = X_2(c_1, c_2, \sqrt{\lambda_3}) \text{ as in (22).}\end{aligned}$$

In particular, *in the symmetric case,*

$$(24)\quad 0 < -x_{1,n} = x_{n,n} < \max\left(\frac{\sqrt{\lambda_3} + \sqrt{\lambda_3 + 4\lambda_2}}{2}, \sqrt{\lambda_i} + \sqrt{\lambda_{i+1}}, 2\sqrt{\lambda_n}\right) \quad (i = 2, 3, \cdots, n-1).$$

We thus obtained bounds for the zeros of $\Phi_n(x)$ *applicable to any sequence of OP.* Leaving aside the discussion of the best possible choice of the constants k_i, k'_i in (22), we clearly see that (23) proves (i).[6] (ii) We have in $(0, \infty)$

$$\lambda_n = b_{2n-2} b_{2n-1},\ c_n = b_{2n-1} + b_{2n} \qquad (n \geqq 2;\ \lambda_1 = b_1;\ c_1 = b_2),$$

where the positive sequence $\{b_n\}$ is unbounded.[7] (iii) This is an immediate consequence of (i, ii). In particular, *in the symmetric case corresponding to* $(-\infty, \infty)$, *the sequence* $\{\lambda_n\}$ *is unbounded.*

Remark. The preassignment of the sequence $\{\lambda_n\}$ determines all Δ_n (see (10, 11)).

4. *Construction of the corresponding sequence of AP.* The differential equations (I, J, L) determine polynomials of degree $n = 0, 1, 2, \cdots$,

---

[6] Cf. our *Mémorial Fasc.*, p. 41, where the same method, applied to the symmetric case only, leads to

$$-x_{1,n} = x_{n,n} < \max(2\sqrt{\lambda_i}) \qquad (i = 1, 2, \cdots, n).$$

(an obvious misprint gives $x_{1,n} > \cdots$). Cf. also: O. Bottema, "Die Nullstellen gewisser durch Recursionformeln definierten Polynome," *Amsterdam Acad. Sc., Proc. Sect. Sc., V*, vol. 34 (1931), pp. 681-691, where the sufficiency of (i) (bounds for the zeros of $\Phi_n(x)$) is established by means of the theory of quadratic forms, also a proof of a part of (iii) is given, different from the one below.

[7] Stieltjes, "Recherches sur les fractions continues," *Oeuvres*, vol. 2, pp. 402-566.

$$
(25)\quad
\begin{aligned}
(J):\ J_n \equiv J_n(x;\alpha,\beta) &= x^n - \frac{n(n+\alpha-1)}{2n+\alpha+\beta-2}\,x^{n-1} \\
&\quad + \frac{n(n-1)(n+\alpha-1)(n+\alpha-2)}{(2n+\alpha+\beta-2)(2n+\alpha+\beta-3)}\,x^{n-2} \mp \cdots \\
(L):\ l_n \equiv l_n(x;\alpha) &= x^n - \frac{n(n+\alpha-1)}{1!}\,x^{n-1} \\
&\quad + \frac{n(n-1)(n+\alpha-1)(n+\alpha-2)}{2!}\,x^{n-2} + \cdots
\end{aligned}
$$

for all α, β, unless a coefficient in $(25, J)$ has a vanishing denominator, while its numerator is different from 0 (if both vanish, we take this coefficient as $= 0$). Furthermore,

$$(26)\qquad J_n(x;\alpha,\beta) \equiv T_n(x;\alpha,\beta)\,;\, l_n(x;\alpha) \equiv L_n(x;\alpha) \qquad (\alpha,\beta > 0).$$

$$(27)\qquad J'_n(x;\alpha,\beta) \equiv nJ_{n-1}(x;\alpha+1,\beta+1)\,;\, l'_n(x;\alpha) \equiv nl_{n-1}(x;\alpha+1) \qquad (\text{for all } \alpha,\beta).$$

Assuming again $\alpha, \beta > 0$, we start with a certain $J_n(x;\alpha,\beta)$ or $l_n(x;\alpha)$ and construct correspondingly infinite sequences of polynomials

$$
(28)\quad
\begin{aligned}
&(J):\ J_0, J_1(x;\alpha+n-1,\beta+n-1),\cdots, J_n(x;\alpha,\beta), \\
&\qquad\qquad J_{n+1}(x;\alpha-1,\beta-1),\cdots \\
&(L):\ l_0, l_1(x;\alpha+n-1),\cdots, l_n(x;\alpha), l_{n+1}(x;\alpha-1),\cdots
\end{aligned}
$$

of degrees $0, 1, 2, \cdots$. It is readily seen, from (25), that they are all well determined. *By virtue of* (27), *each sequence* (28) *is an AP sequence.* For Hermite polynomials we have an *AP* sequence without any additional construction [see (5)]

$$(28\text{ bis})\qquad H_0, H_1(x),\cdots, H_n(x), H_{n+1}(x),\cdots.$$

We now proceed to find the *generating functions* for (28). We recall from the theory of *AP* that the formal power series expansion

$$(29)\qquad a(h) = \gamma_0 + \gamma_1\frac{h}{1!} + \cdots + \gamma_n\frac{h^n}{n!} + \cdots$$

gives rise to an *AP* sequence $\{A_n(x)\}$ through the expansion

$$(30)\qquad a(h)e^{hx} \equiv A_0 + A_1(x)\frac{h}{1!} + \cdots + A_n(x)\frac{h^n}{n!} + \cdots.$$

Conversely, given a sequence of constants $\gamma_0(\neq 0), \gamma_1, \cdots, \gamma_n, \cdots$, constructs

$$(31)\qquad A_n(x) = \gamma_0 x^n + \binom{n}{1}\gamma_1 x^{n-1} + \binom{n}{2}\gamma_2 x^{n-2} + \cdots + \gamma_n \qquad (n = 0, 1, 2, \cdots).$$

The sequence $\{A_n(x)\}$ is then an AP sequence, and $a(h)$, defined by the expansion (29), (assumed to converge for $|h|$ sufficiently small), is said to be its generating function. We may write (31) symbolically (this will be designated by $\simeq$) as

$$(32)\qquad A_n(x) \simeq (x+\bar{\gamma})^n \qquad (n=1,2,\cdots;\ A_0=\gamma_0)$$

where $\bar{\gamma}$ means that we agree to replace each γ^i by γ_i, $(i=0,1,\cdots,n)$.

Regarding the sequences (28), we read at once from (25)

$$(J):\ \gamma_0=1,\ \gamma_1=-1!\cdot\frac{n+\alpha-1}{2n+\alpha+\beta-2}$$

$$\gamma_2=+2!\,\frac{n(n-1)(n+\alpha-1)(n+\alpha-2)}{(2n+\alpha+\beta-2)(2n+\alpha+\beta-3)},\cdots$$

$$(L):\ \gamma_0=1,\ \gamma_1=-(n+\alpha-1),\ \gamma_2=+(n+\alpha-1)(n+\alpha-2),\cdots$$

whence,

$$(33)\qquad \begin{aligned}&(J):\ a(h)=G(-n-\alpha+1,-2n-\alpha-\beta+2,-h)\\ &\qquad\left(G(\alpha,\beta,x)\equiv 1+\frac{\alpha}{1\cdot\beta}x+\frac{\alpha(\alpha+1)}{1\cdot 2\beta(\beta+1)}x^2+\cdots\right)\\ &(L):\ a(h)=(1-\mathrm{h})^{n+\alpha-1}.\end{aligned}$$ [8]

A direct application of (33) is the expression of x^n in terms of the polynomials (25). By the general theory of AP, if the sequence $\{B_n(x)\}$ is generated by $1/a(h)\equiv\beta_0+\beta_1h/1!+\beta_2h^2/2!+\cdots$, then

$$(34)\qquad x^n\simeq(\bar{A}+\bar{\beta})^n\quad(\simeq(AB)_n\equiv(BA)_n),\ \text{where } B_n(x)\simeq(x+\bar{\beta})^n$$

[the symbolical expressions $(\bar{A}+\bar{\beta})^n$, $(x+\bar{\beta})^n$ being interpreted as $(x+\bar{\gamma})^n$ is in (32)]. The application to Laguerre polynomials is particularly simple, since here $1/a(h)=(1-h)^{-(n+\alpha-1)}$. Thus,

$$(35)\qquad x^n=\sum_{i=0}^{n}(-1)^i\cdot i!\cdot\binom{-(n+\alpha-1)}{i}L_{n-i}(x;\alpha+i).$$

Consider (34) for the special case

$$(36)\qquad 1/a(h)\equiv a(-h).$$

A glance at (29, 31) shows that here

$$\beta_i=(-1)^i\gamma_i\ (i=0,1,\cdots);\qquad B_n(x)\simeq(x-\bar{\gamma})^n,$$

and the following very simple reciprocal relation results

$$(37)\qquad A_n(x)\simeq(x+\bar{\gamma})^n,\qquad x^n\simeq(\bar{A}-\bar{\gamma})^n\qquad(n=1,2,\cdots).$$

[8] Cf. Appell, *loc. cit.*, where (33, L) is derived, without, however, indicating its relation to Laguerre polynomials.

This is the case of Hermite polynomials.[9]

5. *The sequences* (25) *and orthogonality.* In (25) each term, by virtue of (26), belongs to a *distinct* sequence of *OP*, as long as the parameters $\alpha \pm \nu$, $\beta \pm \nu$ remain positive. We ask now: can $(25, J)$ and or $(25, L)$, *taken as a whole*, form a sequence of *OP*? That this is conceivable follows from the fact that if we take in (7) c_i, λ_i $(i = 1, 2, \cdots, n)$ from one "permissible" sequence (i. e. all $\lambda_n > 0$), and $c_{n+1}, \lambda_{n+1}, \cdots$ from another such sequence, the first n *OP* thus derived belong to *two* sequences of *OP*.

The answer to our question is based upon Theorem I and does not utilize the properties of the zeros of the polynomials under discussion, which will be denoted by

$$A_0, A_1, \cdots, A_n, \cdots \qquad (A_n = x^n + \cdots).$$

Making use of (6), rewrite the differential equation (1) as

$$A_n = -\frac{nB}{C_n} A_{n-1} - \frac{n(n-1)A}{C_n} A_{n-2}. \tag{38}$$

$$((J): B \equiv B(\alpha, \beta), C_n \equiv C_n(\alpha, \beta); \quad (L): B \equiv B(\alpha), C_n \equiv C_n(\alpha)).$$

If $\{A_n\}$ is a sequence of *OP*, then

$$A_n = (x - c_n) A_{n-1} - \lambda_n A_{n-2} \qquad (c_n, \lambda_n \text{ are constants}),$$

whence,

$$\left(x - c_n + \frac{nB}{C_n}\right) A_{n-1} = \left(\lambda_n - \frac{n(n-1)A}{C_n}\right) A_{n-2}.$$

But two successive *OP*'s cannot have common zeros; hence, necessarily,

$$x - c_n + \frac{nB}{C_n} \equiv 0; \quad \lambda_n - \frac{n(n-1)A}{C_n} \equiv 0, \quad \text{i. e.} \quad A \equiv \text{const.},$$

which holds for Hermite polynomials only.

We indicate in passing some relations for A, B, C in (1). Differentiating (37) and using (6) once more, we have

$$A_{n-1}(C_n + B') + (n-1) A_{n-1}(A' + B) + (n-1) A A_{n-3} = 0.$$

On the other hand, by (38), changing n into $n-1$ and *increasing* α, β *by* 1, (which is indicated by writing $\bar{B}$, $\bar{C}_{n-1}$), we have

[9] We have:

$$H_n \equiv H_n(x; e^{-x^2/4})$$
$$= x^n - \frac{n(n-1)}{1!} x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!} x^{n-4} \mp \cdots; \quad a(h) = e^{-h^2},$$

so that

$$x^n = H_n + \frac{n(n-1)}{1!} x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!} x^{n-2} + \cdots.$$

$$A_{n-1}\bar{C}_{n-1} + (n-1)A_{n-2}\bar{B} + (n-1)(n-2)AA_{n-3} = 0.$$

Hence, the following relations must be satisfied for all COP:

$$C_n + B' - \bar{C}_{n-1} = 0, \qquad A' + B. - \bar{B} = 0.$$

We see once more, that for Hermite polynomials, where no parameters are involved, so that $\bar{B} \equiv B$, we must have $A' = 0$, $A \equiv \text{const.}$

6. *The sequence $\{H_n(x)\}$ as an AP sequence.* We close our discussion by proving a very general

THEOREM IV. *The only system of $OP\{\Phi_n(x; d\psi)\}$ which is at the same time an AP sequence is that with $d\psi(x) = e^{-h^2(x-c)^2}$ (h, c-const.), i. e. that which is reducible to Hermite polynomials by a linear transformation.*[10]

Proof. Assume that

$$\Phi'_n(x; d\psi) \equiv n\Phi_{n-1}(x; d\psi) \qquad (n = 1, 2, \cdots). \tag{39}$$

Combining (39) with the relation obtained by differentiating (7), we get, with the notations (3),

$$\frac{S_n}{n} = \frac{S_{n-1}}{n-1}; \qquad n\lambda_{n-1} - (n-2)\lambda_n = -\frac{2d_{n-1,n-3}}{n-1}. \tag{40}$$

Combining this with

$$c_n = S_n - S_{n-1} \qquad (\text{see } (17))$$

leads to the fundamental result:

$$c_n = \frac{S_n}{n} = \frac{S_{n-1}}{n-1} = c_{n-1}, \quad \text{i. e.} \quad c_n = c_{n-1} = \cdots = c_1 = \text{constant } c.$$

By the remark made above, *the linear substitution $(x - c)\,|\,x$ reduces the system $\{\Phi_n(x; d\psi)\}$ under discussion to the symmetric case.*

Assuming this reduction to have been made and keeping the old notations, we have now

$$c_n = S_n = 0, \qquad \alpha_{2n-1} = 0 \qquad (n = 1, 2, \cdots). \tag{41}$$

---

[10] Meixner, "Orthogonale Polynome mit einer besonderen Gestalt der erzeugenden Funktion," *Journal of the London Mathematical Society*, vol. 9 (1934), pp. 6-13, derives a similar theorem for polynomials satisfying a more general relation than (6), by considerations different from those developed below. A different proof of Theorem IV is to be found in the Thesis of my pupil Dr. M. Webster. Our proof is a straightforward one, applicable to similar problems. Cf. also the interesting article of I. M. Sheffer, "A differential equation for Appell polynomials," *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 914-923.

Making use of the expression of $\Phi_n(x)$ in terms of the moments α_n, we rewrite (39) as follows:

$$(42)\qquad \begin{vmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_{n+1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \alpha_{n-1} & \alpha_n & \cdots & \alpha_{2n-1} \\ 0 & 1 & x \cdots & nx^{n-1} \end{vmatrix} : \Delta_n = n \begin{vmatrix} \alpha_0 & \alpha_1 \cdots & \alpha_{n-1} \\ \alpha_1 & \alpha_2 \cdots & \alpha_n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \alpha_{n-2} & \cdots & \alpha_{2n-3} \\ 1 & x \cdots & x^{n-1} \end{vmatrix} : \Delta_{n-1}$$

$$(n = 2, 3, \cdots).$$

Writing (42) for $n = 3$ and making use of (41, 11), we get successively

$$(43)\qquad \alpha_4 = 3\alpha_0\gamma^2, \quad \alpha_6 = 15\alpha_0\gamma^3, \quad \text{where} \quad \alpha_2/\alpha_0 = \gamma > 0,$$

$$(44)\qquad \Delta_2 = \alpha_0^2\gamma, \quad \Delta_3 = 2\alpha_0^3\gamma^3, \quad \Delta_4 = 24\alpha_0^4\gamma^6,$$

$$(45)\qquad \lambda_2 = \gamma, \qquad \lambda_3 = 2\gamma, \qquad \lambda_4 = 3\gamma.$$

It remains to prove that

$$(46)\qquad \lambda_n = (n-1)\gamma \quad \textit{for all } n.$$

Here we use mathematical induction in the following manner: The recurrence relation (7) shows in the symmetric case, by comparing coefficients, that

$$d_{n,n-2} = -(\lambda_2 + \lambda_3 + \cdots + \lambda_n),$$

so that the second relation (40) gives, assuming that (46) holds up to λ_{n-1} inclusive,

$$n(n-2)\gamma - (n-2)\lambda_n = (n-2)\gamma; \qquad \lambda_n = (n-1)\gamma.$$

The induction is complete, and Theorem IV is established, for the sequence $\{c_n = 0,\ \lambda_n = (n-1)\gamma\}$ gives rise to the system of Hermite polynomials $H_n(x; e^{-x^2/\gamma})$.

THE UNIVERSITY OF PENNSYLVANIA.

ANALYTIC FUNCTIONS STAR-LIKE IN ONE DIRECTION.

By M. S. Robertson.[1]

Introduction. Let (S) denote the class of functions $f(z)$ which have one or the other of the following sequences of properties:

Either (A): 1. $f(z) = z + \sum_2^\infty a_n z^n$ is regular for $|z| < 1$.

2. There exists a positive $\delta = \delta(f)$ so that for every r in the open interval $1 - \delta < r < 1$ $f(z)$ maps $|z| = r$ into a contour C_r which is cut by the real axis in two, and not more than two, points.

or (B): 1. $f(z) = z + \sum_2^\infty a_n z^n$ is regular for $|z| \leqq 1$.

2. $f(z)$ maps $|z| = 1$ into a contour which is cut by the real axis in two, and not more than two, points.

If $f(z)$ is a member of (S) we shall say that it is star-like in the direction of the real axis with respect to the unit circle. A function $f(z)$ which is star-like in a direction other than that of the real axis can be reduced to a function of the type considered above by taking $e^{i\alpha}f(e^{-i\alpha}z)$ with a suitable choice for the real parameter α. If $f(z)$ is a member of (S) satisfying (A) 2 there will exist two, and only two points $z_1 = re^{i\theta_1(r;f)}$ and $z_2 = re^{i\theta_2(r;f)}$ at which the imaginary part of $f(z)$, or $If(z)$, is zero. Moreover, if

$$f(re^{i\theta}) = u(r,\theta) + iv(r,\theta)$$

then

$$(1.1) \qquad \begin{aligned} v(r,\theta) &> 0 \quad \text{when} \quad \theta_1(r) < \theta < \theta_2(r) \\ v(r,\theta) &< 0 \quad \text{when} \quad \theta_2(r) < \theta < \theta_1(r) + 2\pi. \end{aligned}$$

We may define $\theta_1(r) = \theta_1(r;f)$ and $\theta_2(r) = \theta_2(r;f)$ so that

$$(1.2) \qquad 0 \leqq \theta_1(r) \leqq 2\pi, \qquad 0 < \theta_2(r) - \theta_1(r) < 2\pi.$$

When $f(z)$ is real on the real axis (a_n real) then $f(z)$ belongs to the sub-class of (S), consisting of typically-real functions defined by W. Rogosinski.[2] In this case,

[1] National Research Fellow.

[2] See W. Rogosinski, "Über positive harmonische Entwicklungen und typisch-reelle Potenzreihen," *Mathematische Zeitschrift*, Band 35 (1932), pp. 93-121.

$$\theta_1(r) \equiv 0, \qquad \theta_2(r) \equiv \pi.$$

We shall show in this paper that if $f(z)$ is a member of (S) then the coefficients a_n of $f(z)$ satisfy the inequalities

$$|a_n| \leqq n^2 \quad \text{for all } n$$

and the equality is attained for any fixed n by only one function of class (S), namely $(z+\epsilon z^2)/(1-\epsilon z)^3$, $(\epsilon = \pm i)$. If further, $f(z)$ is an odd function, then

$$|a_n| \leqq n \quad \text{for all } n.$$

and the equality sign is attained by the function $(z+z^3)/(1-z^2)^2$. If $f(z)$ is not necessarily odd but is real on the real axis then

$$|a_n| \leqq n \quad \text{for all } n,$$

and the equality sign is atttained by $z/(1-z)^2$. We shall show also that if $f(z)$ belongs to (S) $\lim\limits_{r \to 1} f(re^{i\theta})$ exists and is finite for almost all θ.

If $g(z) = z + \sum\limits_2^\infty c_n z^n$ is regular and univalent for $|z| < 1$ then it has been conjectured[3] that $|c_n| \leqq n$. This is known to be true in case $g(z)$ is star-like in the unit circle, or when $R\sqrt{g(z)/z} \geqq \frac{1}{2}$ for $|z| < 1$, or when there exists a complex constant α so that $R\alpha z g'(z)/g(z) \geqq 0$ for $|z| < 1$, or, finally, when $g(z)$ is real on the real axis. We shall show here that the conjecture is also true when $g(z)$ has the property that it maps $|z| = r < 1$ for every r near 1 into a contour which is convex in one direction, i. e., every straight line parallel to this direction may cut the contour in not more than two points. The equality sign is attained for any fixed n for essentially only one function of this class,

$$z/(1-e^{i\alpha}z)^2, \quad (\alpha \text{ real}).$$

2. *A representation for functions of class* (S). Suppose $f(z)$ is a member of (S) satisfying A (2) and hence (1.1) and (1.2). Let

$$\phi(r) \equiv \frac{\pi}{2} - \frac{\theta_2(r) + \theta_1(r)}{2}, \qquad 1-\delta < r < 1,$$

$$F_r(z) \equiv f(re^{-i\phi(r)}z)\left[\frac{1}{z} - z + 2i\sin\left\{\frac{\pi}{2} - \frac{\theta_2(r) - \theta_1(r)}{2}\right\}\right]$$

If $$f(re^{-i\phi(r)}z) = P(r,\theta) + iQ(r,\theta), \qquad z = e^{i\theta},$$

[3] See L. Bieberbach, *Sitzber. kgl. Akad. Berlin* 1916, pp. 940-955.

then from (1.1)

$$
(2.1)\quad \begin{aligned} &Q(r,\theta) > 0 \text{ when } \frac{\pi}{2} - \frac{\{\theta_2(r)-\theta_1(r)\}}{2} < \theta < \frac{\pi}{2} + \frac{\{\theta_2(r)-\theta_1(r)\}}{2} \\ &Q(r,\theta) < 0 \text{ when } \frac{\pi}{2} + \frac{\{\theta_2(r)-\theta_1(r)\}}{2} < \theta < \frac{5\pi}{2} - \frac{\{\theta_2(r)-\theta_1(r)\}}{2} \end{aligned}
$$

$$
F_r(e^{i\theta}) = \{P(r,\theta) + iQ(r,\theta)\}\left[-2i\sin\theta + 2i\sin\left(\frac{\pi}{2} - \frac{\theta_2(r)-\theta_1(r)}{2}\right)\right].
$$

$$
\mathcal{R}F_r(e^{i\theta}) = 2Q(r,\theta)\left[\sin\theta - \sin\left(\frac{\pi}{2} - \frac{\theta_2(r)-\theta_1(r)}{2}\right)\right]
$$

$$
\geqq 0 \text{ for all } \theta \text{ by (2.1).}
$$

Hence $\mathcal{R}F_r(z) \geqq 0$ for $|z| = 1$; $F_r(z)$ is regular on $|z| = 1$, $F_r(0) \neq 0$; and since the minimum of a harmonic function occurs on the boundary we have

$$
(2.2)\qquad \mathcal{R}F_r(z) \geqq 0 \quad \text{for} \quad |z| \leqq 1.
$$

Let $\{r_i\}$ be a sequence of values of r tending to 1 so that

$$
(2.3)\qquad \lim_{r_i\to 1}\theta_1(r_i) = \alpha, \qquad \lim_{r_i\to 1}\theta_2(r_i) = \beta.
$$

On account of (1.2) we have

$$
0 \leqq \alpha \leqq 2\pi, \qquad 0 \leqq \beta - \alpha \leqq 2\pi.
$$

Let $\mu = \dfrac{\alpha+\beta}{2} - 2\kappa\pi$ with κ so chosen that $0 \leqq \mu < 2\pi$. Let $\nu = \dfrac{\beta-\alpha}{2}$.

Then $0 \leqq \nu \leqq \pi$. Let

$$
F_1(z) \equiv \lim_{r_i\to 1} F_{r_i}(z) = f(-ie^{i\mu}z)\cdot\left(\frac{1+2iz\cos\nu - z^2}{z}\right)
$$

$F_1(z)$ is regular for $|z| < 1$, and by (2.2)

$$
\mathcal{R}F_1(z) \geqq 0 \text{ for } |z| < 1,\ F_1(0) = \sin\mu - i\cos\mu, \text{ whence } \sin\mu \geqq 0,
$$

equality holding only when $F_1(z)$ reduces to either $+i$ or $-i$. In this case $f(z)$ has the form $z(1-2z\cos\nu + z^2)^{-1}$. In any case we have $0 \leqq \mu \leqq \pi$, $0 \leqq \nu \leqq \pi$. If $\sin\mu \neq 0$ we let

$$
F(z) \equiv (\sin\mu)^{-1}\cdot[F_1(ie^{-i\mu}z) + i\cos\mu].
$$

If $\sin\mu = 0$, we may take $F(z) \equiv 1$. In both cases $F(z)$ is regular for $|z| < 1$, $F(0) = 1$, $\mathcal{R}F(z) > 0$ for $|z| < 1$, and

$$
(2.4)\qquad f(z) = h_\nu(e^{-i\mu}z)\cdot(\cos\mu + i\sin\mu\cdot F(z))
$$

where

$$
h_\nu(z) \equiv z(1-2z\cos\nu + z^2)^{-1}, \qquad 0 \leqq \mu \leqq \pi.
$$

Since $\Re F(z) \geqq 0$ the function $\{1 + F(z)\}^{-1}$ is bounded in the unit circle, and so tends to a limit, different from zero, for almost all values of θ as $r \to 1$ $(z = re^{i\theta})$. Hence $\lim_{r\to 1} F(re^{i\theta})$ exists, and is finite for almost all values of θ. Therefore $\lim_{r\to 1} f(re^{i\theta})$ exists, and is finite for almost all values of θ.

3. *Univalent functions convex in one direction.* Let

$$g(z) = z + \sum_{2}^{\infty} c_n z^n \tag{3.1}$$

be a member of a class $(\mathfrak{F})$ of functions regular for $|z| < 1$ which are univalent and convex in one direction. Without loss of generality we may assume that this is the direction of the imaginary axis. We suppose then that there is a positive $\delta = \delta(g)$ so that for every r in the open interval $1 - \delta < r < 1$ $g(z)$ maps $|z| = r$ into a contour C_r such that every straight line parallel to the imaginary axis cuts C_r in not more than two points. If $g(z)$ is regular on $|z| = 1$ we may take $\delta = 0$ in our definition. It is readily seen that a necessary and sufficient condition that $g(z)$ map each circle $|z| = r$ on contours C_r which are of the type described above is that for every r in the interval $1 - \delta < r < 1$ there should exist two real numbers $\theta_1(r)$ and $\theta_2(r)$ satisfying (1.2) so that $\Re g(re^{i\theta})$ is a monotone decreasing function of θ in the interval (θ_1, θ_2) and a monotone increasing function of θ in the complementary interval $(\theta_2, \theta_1 + 2\pi)$. But since

$$I\{zg'(z)\} = -\frac{\partial \Re g(re^{i\theta})}{\partial \theta}$$

it follows that $g(z)$ is a member of class $(\mathfrak{F})$ if, and only if, $zg'(z)$ belongs to class (S). If $zg'(z)$ belongs to (S) then $g(z)$ is univalent. For if not, let

$$g(z_1) = g(z_2) = w_0$$

for two points z_1 and z_2 $(z_1 \neq z_2)$ lying within $|z| = r$ where r is in the interval $1 - \delta < r < 1$. Then as z describes the circle $|z| = r$ enclosing z_1 and z_2 $g(z)$ describes a continuous closed contour C_r which must consist of at least two loops [4] about the point w_0. If this were the case a straight line parallel to the imaginary axis would cut C_r in more than two points, which is a contradiction. Hence $g(z)$ is univalent for $|z| < 1$.

We may then represent $g(z)$ in the form

$$g(z) = \int_0^z \frac{1 + ie^{-i\mu} \sin \mu\{F(z) - 1\}}{1 - 2ze^{-i\mu} \cos \nu + e^{-2i\mu} z^2}\, dz \tag{3.2}$$

[4] See for example Titchmarsh, *The Theory of Functions*, 1932, p. 201.

where $\mathcal{R}F(z) \geqq 0$ for $|z| < 1$, $0 \leqq \mu \leqq \pi$, $0 \leqq \nu \leqq \pi$. Conversely, if $F(z)$ is an arbitrary function with positive real part and regular for $|z| < 1$, and μ any parameter satisfying the inequalities $0 \leqq \mu \leqq \pi$, then the function $g(z)$ formed by the expression (3.2) is univalent for $|z| < 1$ and is the limit of functions $g_n(z)$ each of which is regular and univalent for $|z| < 1$, has at most either a simple pole or two logarithmic singularities on $|z| = 1$, and which map $|z| = 1$ on a contour convex in the direction of the imaginary axis. For let $\{r_n\}$ be a sequence of values of tending to 1 as $n \to \infty$ and define $g_n(z)$ by the equation

$$(3.3) \quad g_n(z) = \int_0^z \frac{1 + ie^{-i\mu} \sin \mu \{F(r_n z) - 1\}}{1 - 2ze^{-i\mu} \cos \nu + e^{-2i\mu} z^2} dz$$

$$(3.4) \quad g(z) = \lim_{n \to \infty} g_n(z) \text{ uniformly in any region interior to the unit circle.}$$

$F(r_n z)$ is regular for $|z| \leqq 1$, $zg'_n(z)$ is regular on $|z| = 1$ save for either one pole of multiplicity 2 ($\nu = 0$) or two simple poles. From § 2 we see that $zg'_n(z)$ maps $|z| = 1$ into a contour which is cut by the real axis in not more than two points. Hence $g_n(z)$ maps $|z| = 1$ into a contour convex in the direction of the imaginary axis and $g_n(z)$ is univalent for $|z| < 1$. It follows by the theorem of Montel [5] that $g(z)$ is also univalent for $|z| < 1$.

If in particular $g(z)$ is also real on the real axis, then $\mu = \pi/2$, $\nu = \pi/2$ and (3.2) takes the form

$$(3.5) \qquad g(z) = \int_0^z \frac{F(z)}{1 - z^2} dz.$$

We may employ here the Stieltjes formula due to Herglotz [6]

$$(3.6) \qquad F(z) = \frac{1}{2\pi} \int_0^{2\pi} \frac{1 + ze^{-i\theta}}{1 - ze^{-i\theta}} d\alpha(\theta)$$

where $\alpha(\theta)$ is non-decreasing in $(0, 2\pi)$. (3.5) can then be written in the form

$$(3.7) \qquad g(z) = \frac{1}{2\pi i} \int_0^{\pi} \log \left\{ \frac{1 - ze^{-i\theta}}{1 - ze^{i\theta}} \right\} \frac{d\alpha(\theta)}{\sin \theta}$$

when the fact is used that $g(z)$ and $F(z)$ are real on the real axis and an integration is performed.

4. *The coefficients of functions of classes* (S) *and* ($\mathfrak{F}$). If in the representation (2.4)

$$F(z) = 1 + \sum_{n=1}^{\infty} b_n z^n$$

[5] See P. Montel, *Bulletin de la Société Mathématique de France*, t. 53 (1925), p. 253.

[6] See G. Herglotz, *Leipziger Berichte*, vol. 63 (1911), pp. 501-511.

since $\mathcal{R}F(z) \geqq 0$ for $|z| < 1$ we have $|b_n| \leqq 2$ for all n. If

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n$$

a comparison of coefficients on both sides of equation (2. 4) yields

$$a_n = \frac{\sin n\nu}{\sin \nu} e^{-(n-1)i\mu} + i \sin \mu \cdot \sum_{k=1}^{n-1} \frac{\sin k\nu}{\sin \nu} \cdot e^{-k\mu i} b_{n-k}. \tag{4. 1}$$

$$|a_n| \leqq \left| \frac{\sin n\nu}{\sin \nu} \right| + 2 \sin \mu \left[1 + \sum_{k=2}^{n-1} \left| \frac{\sin k\nu}{\sin \nu} \right| \right] \tag{4. 2}$$

where $\sin \mu \geqq 0$, $\sin \nu \geqq 0$.

$$|a_n| \leqq n + 2 \sin \mu \cdot \frac{n(n-1)}{2} = n(1 - \sin \mu) + n^2 \sin \mu \tag{4. 3}$$

$$|a_n| \leqq n^2 \text{ for all } n. \tag{4. 4}$$

The equality sign $|a_n| = n^2$ is attained for a fixed n only when $\nu = 0$ or π, $\mu = \pi/2$. In this case

$$f(z) = \frac{zF(z)}{(1 - \epsilon z)^2}, \qquad (\epsilon = \pm i)$$

$$n^2 = |a_n| = |b_{n-1} + 2\epsilon b_{n-2} + \cdots + k\epsilon^{k-1} b_{n-k} + \cdots + n\epsilon^{n-1}|.$$

Equality can occur only when $b_k = 2\epsilon^k$ $(k = 1, 2, \cdots, n-1)$. In this case $f(z)$ must have the form

$$f(z) = \sum_{k=1}^{n} k^2 \epsilon^{k-1} z^k + a_{n+1} z^{n+1} + \cdots. \tag{4. 5}$$

However, since

$$g(z) = \int_0^z \frac{f(z)}{z} dz = z + c_2 z^2 + \cdots$$

is univalent and $|c_2| = 2$, $g(z)$ must be of the form

$$g(z) = z/(1 - e^{i\alpha} z)^2, \quad \alpha \text{ real}$$

as L. Bieberbach has shown.[7]
Hence (4. 5) reduces to

$$f(z) = \frac{z + \epsilon z^2}{(1 - \epsilon z)^3}, \qquad (\epsilon = \pm i). \tag{4. 6}$$

Hence $|a_n| = n^2$ for a fixed n only for the function (4. 6).

Again, if $f(z)$ is any odd function of class (S) then $\beta - \alpha = \pi$, $\nu = \pi/2$. Consequently, from (4. 2) we have

[7] See L. Bieberbach, *loc. cit.*

(4.7) $$| a_{2n+1} | \leqq 1 + 2n \sin \mu \leqq 2n + 1.$$

If $\nu \neq 0$ or π, we also have from (4.2)

$$\overline{\lim_{n\to\infty}} \left| \frac{a_n}{n} \right| \leqq 2 \frac{\sin \mu}{\sin \nu} \cdot M(\nu)$$

where

$$M(\nu) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} | \sin k\nu |.$$

It will follow from the following lemma, due to T. Gronwall,[8] that

(4.8) $$\overline{\lim_{n\to\infty}} \left| \frac{a_n}{n} \right| \leqq \frac{4}{\pi} \cdot \frac{\sin \mu}{\sin \nu}.$$

LEMMA. *If*

$$M_n(\theta) = \frac{1}{n} \sum_{k=1}^{n} | \sin k\theta |$$

then

$$M(\theta) = \lim_{n\to\infty} M_n(\theta)$$

exists and

$$M(\theta) = 2/\pi$$

if θ/π is irrational;

$$M\left(\frac{k}{r}\pi\right) = \frac{\cot \pi/2r}{r} < 2/\pi, \; k \text{ and } r \; (k < r)$$

positive integers prime to each other.

If

$$g(z) = z + \sum_{n=2}^{\infty} c_n z^n$$

is a member of class $(\mathfrak{F})$ then, as we have seen, $zg'(z)$ belongs to (S). Hence

(4.9) $$| c_n | \leqq n \text{ for all } n$$

and equality is attained for any fixed n only by the one function of class $(\mathfrak{F})$, namely

$$z(1 - \epsilon z)^{-2}, \qquad (\epsilon = \pm i).$$

For this particular class of univalent functions this constitutes a proof of the Bieberbach conjecture for the coefficients of a univalent function.

---

[8] See T. Gronwall, "On a theorem of Fejér's," *Transactions of the American Mathematical Society*, 1912, pp. 445-468.

If $g(z)$ is an odd function of class $(\mathfrak{F})$ with coefficients c_{2n+1} then from (4.7) and the fact that $zg'(z)$ belongs to (S) we have

$$(4.10) \qquad | c_{2n+1} | \leqq 1.$$

All the members of class $(\mathfrak{F})$ which are real on the real axis and convex in the direction of the imaginary axis (a direction perpendicular to the line on which the coefficients lie) are given by (3.7) and conversely. The coefficients c_n of $g(z)$ in this case are then given by the formula

$$(4.11) \qquad c_n = \frac{1}{n\pi} \int_0^{\pi} \frac{\sin n\theta}{\sin \theta} d\alpha(\theta)$$

whence $| c_n | \leqq 1$ and equality is attained by

$$z/(1 \pm z), \qquad z/(1 - z^2).$$

THE UNIVERSITY OF CHICAGO.

ON NON-HOMOGENEOUS LINEAR DIFFERENTIAL EQUATIONS OF INFINITE ORDER WITH CONSTANT COEFFICIENTS.

By R. D. Carmichael.

Introduction. The methods of this paper are suggested by the stimulating memoir of A. Hurwitz [1] in which is proved the theorem of C. Guichard [2] that for every integral function $\phi(x)$ the equation $y(x+1) - y(x) = \phi(x)$ has a solution which is itself an integral function. The method also owes much to the remarkable contributions of S. Pincherle [3] to the theory of functional equations, contributions which now extend over a period of nearly fifty years.

In § 1 a formal solution of the differential equation (3) is given and a sufficient condition is set forth (Theorem I) to ensure that it shall be an actual solution. The remainder of the paper consists mainly in presenting effectively workable hypotheses under which this sufficient condition is certainly realized. The maximum simplicity in the formulation of these hypotheses is attained in Theorem VII of § 8, and this theorem is effectively supplemented by Theorem VIII. Auxiliary classifications of integral functions are indicated in §§ 3 and 4. The former is classic, but the latter seems to be new. It is a classification which appeared to be demanded by the course of the argument; it seems to be of interest for its own sake. Attention is called particularly to the invariant point property of integral functions indicated in this connection in § 4.

1. The first general theorem. Let $F(z)$ and $\phi(x)$ denote two given integral functions, neither of them being identically equal to zero, and write their power series expansions in the forms

$$(1) \qquad F(z) = \sum_{\nu=0}^{\infty} a_\nu z^\nu,$$

$$(2) \qquad \phi(x) = \sum_{\nu=0}^{\infty} s_\nu x^\nu / \nu!.$$

We consider the problem of constructing integral functions $y(x)$ satisfying the linear differential equation

$$(3) \qquad a_0 y + a_1 y' + a_2 y'' + \cdots = \phi(x)$$

[1] A. Hurwitz, *Acta Mathematica*, vol. 20 (1897), pp. 285-312.

[2] C. Guichard, *Ann. Ec. Norm. Sup.* (3), vol. 4 (1887), pp. 361-380.

[3] S. Pincherle, *Acta Mathematica*, vol. 46 (1925), pp. 341-362.

of infinite order with constant coefficients. We shall say that $F(z)$ is the *characteristic function* belonging to equation (3).

Let τ be a given positive number. Let $\{\lambda_\nu\}$ and $\{\mu_\nu\}$ $(\nu = 0, 1, 2, \cdots)$ be two infinite sequences of positive numbers such that $\tau \leqq \mu_\nu \leqq \lambda_\nu$ for every ν. For each particular value of ν let C_ν be a closed contour of finite length encircling the point O and lying in the ring $\mu_\nu \leqq |z| \leqq \lambda_\nu$, and let C_ν pass through no point at which $F(z)$ vanishes. Let T_ν denote the sum of the (convergent) series in the relation

$$T_\nu = \sum_{k=0}^{\infty} |a_k| \lambda_\nu{}^k. \tag{4}$$

It is obvious that T_ν is bounded away from zero.

We introduce the functions

$$P_{\nu,n}(x) = \frac{1}{2\pi i}\int_{C_n} \frac{e^{xz}dz}{z^{\nu+1}F(z)} \tag{5}$$

where ν and n take independently the values of the set $0, 1, 2, \cdots$. For their k-th derivatives with respect to x we have

$$P_{\nu,n}^{(k)}(x) = \frac{1}{2\pi i}\int_{C_n} \frac{z^k e^{xz}dz}{z^{\nu+1}F(z)}. \tag{6}$$

We have

$$\sum_{k=0}^{\infty} a_k P_{\nu,n}^{(k)}(x) = \frac{1}{2\pi i}\int_{C_n} \sum_{k=0}^{\infty} a_k z^k \frac{e^{xz}dz}{z^{\nu+1}F(z)} = \frac{1}{2\pi i}\int_{C_n} \frac{e^{xz}dz}{z^{\nu+1}} = \frac{x^\nu}{\nu!}. \tag{7}$$

From (6) we have

$$|P_{\nu,n}^{(k)}(x)| \leqq \frac{\lambda_n{}^k e^{|x|\lambda_n}}{2\pi\mu_n{}^{\nu+1}}\int_{C_n} \frac{|dz|}{|F(z)|}. \tag{8}$$

Let us define $y(x)$ by the relation

$$y(x) = \sum_{\nu=0}^{\infty} s_\nu P_{\nu,\nu}(x) \equiv \sum_{\nu=0}^{\infty} \frac{s_\nu}{2\pi i}\int_{C_\nu} \frac{e^{xz}dz}{z^{\nu+1}F(z)}, \tag{9}$$

subject to appropriate conditions of convergence to be indicated later.

Using the notation thus described, we shall prove the following general theorem:

THEOREM I. *A sufficient condition that $y(x)$, as defined by (9), shall be an integral function satisfying (3) is that the series*

$$\sum_{\nu=0}^{\infty} e^{\rho\lambda_\nu} |s_\nu| T_\nu \mu_\nu{}^{-\nu-1}\int_{C_\nu} \frac{|dz|}{|F(z)|} \tag{10}$$

shall be convergent for every positive number ρ.

Let S be any preassigned finite closed region of the x-plane. Let ρ be the maximum value of $|x|$ for x in S. Since by hypothesis series (10) is convergent and since T_ν is bounded away from zero, it follows that the series obtained from (10) on replacing T_ν by 1 is convergent. From this fact and from relation (8) with $n=\nu$ and $k=0$ we see that for all x in S the series in (9) is dominated term by term by a convergent series of constants. Hence that series is absolutely and uniformly convergent in S and $y(x)$ is an integral function. Moreover, $y^{(k)}(x)$ may be formed from (9) by term-by-term differentiation.

If we proceed formally we have by aid of (7) the relations

$$\begin{aligned}(11)\qquad \sum_{k=0}^{\infty} a_k y^{(k)}(x) &= \sum_{k=0}^{\infty}\sum_{\nu=0}^{\infty} a_k s_\nu P_{\nu,\nu}^{(k)}(x)\\ &= \sum_{\nu=0}^{\infty} s_\nu \sum_{k=0}^{\infty} a_k P_{\nu,\nu}^{(k)}(x)\\ &= \sum_{\nu=0}^{\infty} s_\nu \frac{x^\nu}{\nu!} = \phi(x).\end{aligned}$$

From (8) it follows that the series in the third member of (11) is dominated term-by-term by the series

$$(12)\qquad \sum_{\nu=0}^{\infty} |s_\nu| \sum_{k=0}^{\infty} |a_k| \lambda_\nu^k e^{|x|\lambda_\nu} \mu_\nu^{-\nu-1} \int_{C_\nu} \frac{|dz|}{|F(z)|}.$$

From the fact that $F(z)$ is an integral function it follows that the series here denoted by the second summation sign is convergent. Employing (4), we may then write the series denoted by the first summation sign in the form

$$(13)\qquad \sum_{\nu=0}^{\infty} e^{|x|\lambda_\nu} |s_\nu| T_\nu \mu_\nu^{-\nu-1} \int_{C_\nu} \frac{|dz|}{|F(z)|}.$$

For all x in S this series is dominated term-by-term by the series in (10). It follows that the repeated series in the third member of (11) is absolutely and uniformly convergent in S. Therefore the repeated series in the second member of (11) is absolutely and uniformly convergent in S and its sum is equal to that of the series in the third member. Thence it follows readily from (11) that $y(x)$ is a solution of (3).

These results imply Theorem I.

Corollary. *The k-th derivative of $y(x)$ is dominated as follows:*

$$(14)\qquad |y^{(k)}(x)| \leqq \sum_{\nu=0}^{\infty} e^{|x|\lambda_\nu} |s_\nu| \lambda_\nu^k \mu_\nu^{-\nu-1} \int_{C_\nu} \frac{|dz|}{|F(z)|}, \qquad (k=0,1,2,\cdots).$$

2. Consequences of Theorem I. We shall now prove the following theorem, retaining our previous notation and employing additional hypotheses:

THEOREM II. *Let $F(z)$ be a function of exponential type. Let λ_ν be further restricted by the condition $\lambda_\nu \leqq \alpha(\nu+1)$ $(\nu = 0, 1, 2, \cdots)$, where α is a given positive constant. Let the defined elements be so related that a positive constant β exists for which the relations*

$$\int_{C_\nu} \frac{|\,dz\,|}{|\,F(z)\,|} < \frac{\mu_\nu{}^{\nu+1}\beta^{\nu+1}}{\nu!}, \qquad (\nu = 0, 1, 2, \cdots), \tag{15}$$

are satisfied. Then the function $y(x)$, defined in (9), is an integral function satisfying equation (3).

It is sufficient to show that the series in (10) is convergent. Since $F(z)$ is of exponential type, say of type q, we have for every positive ϵ a constant K_ϵ such that

$$|\,F(z)\,| \leqq \sum_{\nu=0}^{\infty} |\,a_\nu\,| \cdot |\,z\,|^\nu < K_\epsilon e^{(q+\epsilon)|z|}.$$

Then, since $\lambda_\nu < \alpha(\nu+1)$, we have

$$T_\nu < K_\epsilon e^{\alpha(q+\epsilon)(\nu+1)}.$$

Thence, by aid of (15), we see that series (10) is dominated term-by-term by the series

$$\sum_{\nu=0}^{\infty} K_\epsilon \frac{|\,s_\nu\,|}{\nu!} e^{\rho\alpha(\nu+1)} e^{\alpha(q+\epsilon)(\nu+1)} \beta^{\nu+1}.$$

From (2) and the integral character of $\phi(x)$ it follows that the last foregoing series is convergent for every positive number ρ. Hence the same is true of series (10). Therefore Theorem II is established.

Let a, α, b, c be given constants, $a \leqq \alpha$, and take $\mu_\nu = a(\nu+1)$, $\lambda_\nu = \alpha(\nu+1)$. Then, if for every ν of the set $0, 1, 2, \cdots$ the length of C_ν does not exceed $b^{\nu+1}$ and if $|\,F(z)\,| > c^{-\nu-1}$ for every point z on C_ν, it is easy to show that the hypotheses of Theorem II are satisfied. It is obvious that a suitable constant b exists whenever the paths C_ν are circles about 0 as center and having radii limited by the implied conditions. This yields one of the simplest special cases of the theorem.

For the case when $F(z) = e^z - 1$ we may take the path C_ν to be a circle of radius $\pi(2\nu+1)$. Then it is easy to show (by aid of the periodicity properties of $F(z)$) that a constant M exists such that $|\,F(z)\,| > M$ whenever z

is on a circle C_ν. By taking $\mu_\nu = \lambda_\nu = \pi(2\nu + 1)$ one may readily show that the required conditions are satisfied for the application of Theorem II and hence that the corresponding equation (3) has a solution $y(x)$ which is an integral function. This implies that the equation $y(x+1) - y(x) = \phi(x)$ has a solution $y(x)$ which is an integral function, the proof in this case being essentially that of Hurwitz (*loc. cit.*).

In a similar way it may be shown that the difference equation

$$y(x+n) + c_1 y(x+n-1) + \cdots + c_n y(x) = \phi(x)$$

with constant coefficients has a solution $y(x)$ which is an integral function provided that $\phi(x)$ is an integral function. Here the function $F(z)$ may be written

$$F(z) = e^{nz} + c_1 e^{(n-1)z} + \cdots + c_n = \prod_k (e^z - \rho_k)^{t_k},$$

where the ρ_k are constants and the t_k are positive integers. Suitable contours C_ν may be readily defined such that $F(z)$ is again bounded away from 0 for all z on all C_ν and such that Theorem II becomes applicable. In fact, the same method may be extended,[4] but with increased difficulty, to the more general case of the equation

$$c_0 y(x) + c_1 y(x+\alpha_1) + \cdots + c_m y(x+\alpha_m) = \phi(x)$$

where the c's and α's are constants and $\phi(x)$ is an integral function.

3. Integral functions of class $C(t, q)$. With a view to the extension of Theorem II it is convenient to separate certain integral functions into classes and to note some properties of the several classes.

Let t be a given positive number. Let $g(z)$ be an integral function having the power series expansion

$$g(z) = \sum_{\nu=0}^{\infty} c_\nu z^\nu.$$

We shall say that $g(z)$ is of class $C(t, q)$ if and only if

$$\limsup_{\nu=\infty} |\, (\nu!)^{1/t} c_\nu \,|^{1/\nu} = q < \infty.$$

We shall need the following known theorem[5] concerning these functions:

[4] See R. D. Carmichael, *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 1-28.

[5] This theorem belongs to the classic theory of integral functions of order t and normal type or minimal type. For more general results of a similar character see a forthcoming paper on invariant point properties of integral functions under the joint authorship of R. D. Carmichael, W. T. Martin and M. T. Bird.

Necessary and sufficient conditions in order that an integral function $g(z)$ shall be of class $C(t, q)$ are the following:

(a) numbers σ (zero or positive) shall exist for which it is true that for every positive number ϵ there exists a positive number $N_{\epsilon\sigma}$, depending on ϵ and σ but independent of z, such that for all (finite) values of z we have

$$| g(z) | < N_{\epsilon\sigma}\, e^{(\sigma+\epsilon)^t |z|^t/t};$$

(b) the least possible value of such numbers σ is q.

4. Integral functions of class $K(s)$. Let s be a given positive number not less than 1. Let $\psi(x)$ be an integral function having the power series expansion

$$\psi(x) = \sum_{\nu=0}^{\infty} \sigma_\nu x^\nu.$$

We shall say that $\psi(x)$ is of class $K(s)$ if and only if

$$\lim_{\nu=0} \sigma_\nu^{\nu^{-s}} = 0.$$

It is obvious that the class $K(s_1)$ contains the class $K(s_2)$ if $1 \leqq s_1 < s_2$.

If $\psi(x)$ is of class $K(s)$ then for every positive number ϵ there exists a positive number L_ϵ, independent of ν, such that

$$| \sigma_\nu | < L_\epsilon \epsilon^{\nu^s}.$$

Hence, from the relation

$$\psi^{(n)}(x)/n! = \sum_{\nu=0}^{\infty} \sigma_{\nu+n} \frac{(\nu+n)!}{\nu!\, n!} x^\nu$$

we have

$$| \psi^{(n)}(x)/n! | < L_\epsilon \sum_{\nu=0}^{\infty} \epsilon^{(\nu+n)^s} \frac{(\nu+n)!}{\nu!\, n!} | x |^\nu.$$

Now, since $s \geqq 1$, we have

$$(\nu+n)^s \geqq \nu^s + n^s, \quad \frac{(\nu+n)!}{\nu!\, n!} \leqq (1+\delta)^n (1+\delta^{-1})^\nu,$$

when $n + \nu \geqq 1$, where δ is any positive number whatever. Therefore if we take ϵ less than 1 (as we shall) we have

$$| \psi^{(n)}(x)/n! | < L_\epsilon \epsilon^{n^s} (1+\delta)^n \sum_{\nu=0}^{\infty} \epsilon^{\nu^s} (1+\delta^{-1})^\nu | x |^\nu,$$

whence it follows that

$$\lim_{n=\infty} | \psi^{(n)}(x)/n! |^{n^{-s}} = 0.$$

This property, for $x = 0$, is the defining property for the class $K(s)$. This defining property associated with the point 0 therefore persists for all other finite points. Moreover if it holds for $x = x_0$ it clearly holds also for $x = 0$. *It therefore exhibits an invariant point property of the functions of class $K(s)$.*

We shall now prove the following theorem:

A necessary and sufficient condition that the integral function $\psi(x)$ shall be of class $K(s)$, $s \geqq 1$, is that the series

$$\sum_{\nu=0}^{\infty} | \sigma_\nu | \; | x |^{\nu s}$$

shall converge for every finite value of x.

In the proof of this theorem it is convenient to use the following lemma:

LEMMA. *Let s be a positive number not less than* 1. *Let*

$$u_1 + u_2 + u_3 + \cdots$$

be a series in which $u_k \geqq 0$. Define ρ by the relation

$$\limsup_{\nu=\infty} \, (u_\nu)^{\nu^{-s}} = \rho.$$

Then the given series is convergent if $\rho < 1$ and divergent if $\rho > 1$, while for $\rho = 1$ there is no test.

The usual proof for the classic case $s = 1$ of the lemma holds without essential modification in the general case: it is therefore left to the reader.

From the lemma it follows at once that the series in the theorem converges for all finite x if $\psi(x)$ is of class $K(s)$. On the other hand, if the series converges for all finite x it follows from the lemma that

$$\limsup_{\nu=\infty} \, | \sigma_\nu |^{\nu^{-s}} \, | x | \leqq 1$$

for all finite x, whence we see that $\psi(x)$ is of class $K(s)$. The theorem is therefore established.

5. Generalization of Theorem II. We shall now use the notation of § 1, with the implied hypotheses, in proving the following generalization of Theorem II:

THEOREM III. *Let $F(z)$ be an integral function of class $C(t, q)$. Let $\phi(x)$ be an integral function of class $K(s)$. Let λ_ν be further restricted by the condition $\lambda_\nu \leqq \alpha(\nu+1)^\sigma$ $(\nu = 0, 1, 2, \cdots)$, where α is a given positive constant and σ verifies the relation $1 \leqq \sigma \leqq \min(s, s/t)$. Suppose that $s \geqq t$ and $s \geqq 1$. Let the defined elements be so related that a positive constant β exists for which the relations*

$$\int_{C_\nu} \frac{|dz|}{|F(z)|} < \frac{\mu_\nu{}^{\nu+1}\beta^{(\nu+1)^s}}{\nu!}, \qquad (\nu = 0, 1, 2, \cdots), \tag{16}$$

are satisfied. Then the function $y(x)$, defined in (9), is an integral function satisfying the differential equation (3).

In view of Theorem I it is clearly sufficient to prove that the hypotheses here involved imply the convergence of the series in (10).

From the hypothesis on $F(z)$ and from the theorem in § 3 it follows that for every positive number ϵ there exists a positive number K_ϵ such that

$$|F(z)| \leqq \sum_{\nu=0}^{\infty} |a_\nu| \; |z|^\nu < K_\epsilon e^{(q+\epsilon)^t |z|^t/t}.$$

Then, since $\lambda_\nu < \alpha(\nu+1)^\sigma$, we see from (4) that

$$T_\nu < K_\epsilon \, e^{a^t(q+\epsilon)^t(\nu+1)^s/t}.$$

Thence, by aid of (16), we see that series (10) is dominated term-by-term by the series

$$\sum_{\nu=0}^{\infty} K_\epsilon \frac{|s_\nu|}{\nu!} e^{\rho\alpha(\nu+1)^s} e^{a^t(q+\epsilon)^t(\nu+1)^s/t} \beta^{(\nu+1)^s}.$$

From (2) and from the fact that $\phi(x)$ is of class $K(s)$ and from the theorem of § 4 it follows that the foregoing series is convergent for every positive number ρ. Hence the same is true of series (10). Theorem III is therefore established.

By specialization of the hypotheses in Theorem III we obtain the following theorem which is often more convenient in use than the more general theorem:

THEOREM IV. *Let $F(z)$ and $\phi(x)$ be integral functions of classes $C(t, q)$ and $K(s)$ respectively, where s is not less than t and not less than 1. Let σ be such that $1 \leqq \sigma \leqq \min(s, s/t)$. Let a, α, b, c be given positive constants, $a \leqq \alpha$. If for every number ν of the set $0, 1, 2, \cdots$ there exists a contour C_ν*

encircling the point 0 *and lying in the circular ring* $a(\nu+1) \leqq |z| \leqq \alpha(\nu+1)^\sigma$ *and having a length not exceeding* $b^{(\nu+1)^s}$ *and if* $|F(z)| > c^{-(\nu+1)^s}$ *for every point* z *on* C_ν, *then the function* $y(x)$, *defined in* (9), *is an integral function satisfying equation* (3).

To prove this theorem we show that the hypotheses imply those of Theorem III. We take $\mu_\nu = a(\nu+1)$, $\lambda_\nu = \alpha(\nu+1)^\sigma$. Then, since the hypotheses in Theorem IV imply that

$$\int_{C_\nu} \frac{|dz|}{|F(z)|} < (bc)^{(\nu+1)^s},$$

it is sufficient to observe that a positive number β obviously exists such that

$$(bc)^{(\nu+1)^s} < a^{\nu+1}(\nu+1)^{\nu+1}\beta^{(\nu+1)^s}/\nu!, \qquad (\nu = 0, 1, 2, \cdots),$$

in order to recover all the hypotheses of Theorem III. Hence Theorem IV is established.

6. Another Consequence of Theorem I. We shall now prove the following theorem:

THEOREM V. *Let* $F(z)$ *and* $\phi(x)$ *be integral functions of classes* $C(t, q)$ *and* $K(s)$, *respectively, where* s *is not less than* t *and not less than* 1. *Let* σ *be such that* $1 \leqq \sigma \leqq \min(s, s/t)$. *Let* $C_0, C_1, C_2, \cdots$ *be a sequence of circles about* 0 *as center and of radii* $\sigma_0, \sigma_1, \sigma_2, \cdots$, *respectively, and let them be such that no one of them passes through a point at which* $F(z)$ *vanishes. Moreover, let us suppose that positive constants* a *and* α *exist such that* $a(\nu+1) \leqq \sigma_\nu \leqq \alpha(\nu+1)^\sigma$. *Then, if a positive constant* K *exists such that the relations*

$$(17) \qquad \int_{C_\nu} \frac{|dz|}{|F(z)|} < K^{(\nu+1)^s} \qquad (\nu = 0, 1, 2, \cdots),$$

are satisfied, the function $y(x)$, *defined in* (9) *is an integral function satisfying equation* (3).

We shall show that this proposition is a consequence of Theorem I. We take $\mu_\nu = \lambda_\nu = \sigma_\nu$. Then Theorem I implies the truth of V provided that the series

$$(18) \qquad \sum_{\nu=0}^{\infty} |s_\nu| \, e^{\rho\sigma_\nu} \sigma_\nu^{-\nu-1} \bar{T}_\nu \int_{C_\nu} \frac{|dz|}{|F(z)|}$$

converges for every positive value ρ, the symbol $\bar{T}_\nu$ being defined by the relation

$$\bar{T}_\nu = \sum_{k=0}^{\infty} | a_k | \sigma_\nu{}^k.$$

But series (18) is dominated term by term by the series

$$\sum_{\nu=0}^{\infty} \frac{| s_\nu |}{\nu !} e^{a\rho(\nu+1)^s} a^{-\nu-1} (\nu + 1)^{-\nu-1} \nu ! \bar{T}_\nu K^{(\nu+1)^s}.$$

Now take the $[(\nu+1)^s]$-th root of the ν-th term of this series, noting that for $\bar{T}_\nu$ we have from its definition and the theorem of § 2 a relation of the form

$$\bar{T}_\nu < K_\epsilon \, e^{\alpha^t (q+\epsilon)^t (\nu+1)^s / t}$$

and employing the usual asymptotic form for $\nu !$; as ν becomes infinite this root approaches the limit 0. Hence from the lemma in § 4 it follows that series (18) converges for every positive value of ρ. Therefore the theorem is established.

We shall now show that the following theorem is a corollary of the preceding:

THEOREM VI. *Let $F(z)$ and $\phi(x)$ be integral functions of classes $C(t, q)$ and $K(s)$, respectively, where s is not less than t and not less than* 1. *Let σ be such that $1 \leqq \sigma \leqq \min(s, s/t)$. Let $m(r)$ denote the minimum value of $|F(z)|$ on the circle $|z| = r$. Let $\sigma_0, \sigma_1, \sigma_2, \cdots$ be a sequence of positive numbers such that $a(\nu+1) \leqq \sigma_\nu \leqq \alpha(\nu+1)^\sigma$, where a and α are given positive constants and $a \leqq \alpha$. Let C_ν denote the circle $|z| = \sigma_\nu$. Suppose that a positive number L exists such that the relations*

$$(19) \qquad m(\sigma_\nu) > L^{-(\nu+1)^s}, \qquad (\nu = 0, 1, 2, \cdots),$$

are satisfied. Then the function $y(x)$, defined by (9), *is an integral function satisfying equation* (3).

This theorem is seen to follow from the preceding by observing that conditions (19) imply conditions (18) and also imply that C_ν passes through no point at which $F(z)$ vanishes.

7. Auxiliary properties of integral functions. Let $F(z)$ be an integral function of class $C(t, q)$. In case $F(z)$ has zeros away from the point 0 we denote these by $\alpha_1, \alpha_2, \cdots$ in order of increasing moduli. If the number of zeros is infinite, then it is a classic result that the series

$$\sum_{n=1}^{\infty}\left|\frac{1}{\alpha_n}\right|^{t+\epsilon} \tag{20}$$

converges for every positive number ϵ. Moreover, we have the following well-known theorem:[6]

Let k be any positive number. For each i inclose the zero α_i by a circle with α_i as center and of radius $|\alpha_i|^{-k}$. Then in that part of the finite plane which is exterior to all these circles we have for every preassigned positive ϵ and for $|z|$ sufficiently large, the relation

$$|F(z)| > e^{-|z|^{t+\epsilon}}. \tag{21}$$

Moreover, it is known,[7] and indeed is implied by the convergence of the series in (20), that for every positive ϵ there exists a positive number M_ϵ such that $|\alpha_n|^{-t-\epsilon} < M_\epsilon/n$. If we write $k = k_1(t+\epsilon)$ we then have the existence of $M_{\epsilon k}$ such that $|\alpha_n|^{-k} < M_{\epsilon k}n^{-k_1}$. If we take $k_1 > 1$ (and this we do) and write

$$L_{\epsilon k} = 2M_{\epsilon k}\sum_{n=1}^{\infty} n^{-k_1}$$

we have

$$2\sum_{n=1}^{\infty}|\alpha_n|^{-k} < L_{\epsilon k}.$$

Therefore from the proposition associated with (21) we have the following theorem:

If $F(z)$ is of class $C(t, q)$ and ϵ is a positive number, then a number L exists $(L \geqq L_{\epsilon k})$ such that in every interval of length L sufficiently far out on the positive axis of reals there exists a value r such that

$$|F(z)| > e^{-|z|^{t+\epsilon}} \quad \textit{for} \quad |z| = r. \tag{22}$$

8. Two additional existence theorems for equation (3). We shall now employ the theorem at the end of § 7 in obtaining from Theorem VI the following more simply formulated proposition:

THEOREM VII. *Let $F(z)$ and $\phi(x)$ be integral functions of classes $C(t, q)$ and $K(s)$, respectively, and let t be less than s $(s \geqq 1)$. Then the differential equation (3) has a solution $y(x)$, defined in equation (24) below, such that $y(x)$ is an integral function.*

[6] See *Encyklopädie der Mathematische Wissenschaften,* Band II_3, p. 436.

[7] See *Encyklopädie der Mathematische Wissenschaften,* Band II_3, p. 434.

Let the positive number ϵ be such that $t+\epsilon<s$. Let a and α be two positive numbers such that $a<\alpha$. Then from the result at the end of the preceding section it follows that for sufficiently large values of ν we have on each interval $(a(\nu+1), \alpha(\nu+1))$ a number σ_ν such that

$$|F(z)|>e^{-|z|^{t+\epsilon}}, \qquad |z|=\sigma_\nu. \tag{23}$$

We denote by C_ν the circle $|z|=\sigma_\nu$, defining σ_ν conveniently for values of ν smaller than those involved in (23). We write

$$y(x)=\sum_{\nu=0}^{\infty}\frac{s_\nu}{2\pi i}\int_{C_\nu}\frac{e^{xz}dz}{z^{\nu+1}F(z)}. \tag{24}$$

For σ in Theorem VI we take the value $\sigma=1$. Then the hypotheses of Theorem VI are satisfied when ν is sufficiently large, and hence for all values of ν when σ_ν is suitably defined for an appropriate finite number of values of ν. Therefore Theorem VII is established.

COROLLARY. *If $\phi(x)$ is an integral function and if $F(z)$ is of class $C(t,q)$ with $t<1$, then the differential equation* (3) *has a solution $y(x)$ which is an integral function.*

Whether the corollary holds when $t=1$, and more generally whether Theorem VII holds when $t=s$, I have not been able to determine. That there are cases when $t=s$ and when the conclusion holds is shown at once by the theorems for difference equations cited in the latter part of § 2. These involve cases in which $t=s=1$. But our method does not seem to yield the conclusion of Theorem VII when $t=s$ unless some additional hypotheses are placed on $F(z)$. The result associated with (21) seems not to be sufficiently restrictive for this purpose, since we seem to need an inequality more effective than (22). The nature of the additional information required is indicated by the further hypotheses on $F(z)$ introduced in the following theorem.

THEOREM VIII. *Let $F(z)$ and $\phi(x)$ be integral functions of classes $C(t,q)$ and $K(t)$, respectively, where $t\geqq 1$. Let $\sigma_0, \sigma_1, \sigma_2, \cdots$ be a sequence of positive numbers such that $a(\nu+1)\leqq\sigma_\nu\leqq\alpha(\nu+1)$, where a and α are given positive constants and $a<\alpha$. Let C_ν denote the circle $|z|=\sigma_\nu$. Suppose furthermore that $F(z)$ is such that the σ_ν may be chosen (subject to*

the named conditions) so that a positive constant M and a non-negative constant N exist such that

$$|F(z)| > M e^{-N|z|^t} \tag{25}$$

for all z on the circles C_ν. Then the function $y(x)$, defined by (9), is an integral function satisfying the differential equation (3).

This theorem is an almost immediate corollary of Theorem VI. In the latter theorem we take $s = t$, whence we must have $\sigma = 1$. Then in the two theorems the constants σ_ν belong to the same intervals. Using the symbol $m(r)$ of Theorem VI, we see from (25) that

$$m(\sigma_\nu) > M e^{-N\sigma_\nu{}^t} \geqq M e^{-N\alpha^t(\nu+1)^t} > L^{-(\nu+1)^t},$$

where L is a sufficiently large positive number. Hence we have recovered all the hypotheses of Theorem VI. Therefore Theorem VIII is established.

9. Further properties of the solution of Equation (3). In all the cases treated we have shown that the solution $y(x)$ is an integral function. The corollary to Theorem I gives upper bounds to the values of $|y^{(k)}(x)|$, $(k = 0, 1, 2, \cdots)$, valid in the general case of § 1; and this may obviously be specialized so as to yield more precise inequalities under the additional hypotheses involved in the less general theorems. We shall now show that there is a class of cases, depending on the nature of the characteristic function $F(z)$, in which additional information concerning $y(x)$ may be obtained directly from the convergence of the series in the left member of (3).

Let us suppose that $F(z)$ is of class $C(t, q)$ with q greater than 0 and let the coefficients $a_0, a_1, a_2, \cdots$ have a certain character of regularity implied by the required condition that

$$\lim_{\nu=\infty} |(\nu!)^{1/t} a_\nu|^{1/\nu} = q, \tag{26}$$

the superior limit in the defining condition of the class $C(t, q)$ being thus replaced by an actual limit. Then the condition

$$\limsup_{\nu=\infty} |a_\nu y^{(\nu)}(x)| \leqq 1,$$

necessary for the convergence of the series in the left member of (3), may be written in the form

$$\limsup_{\nu=\infty} |(\nu!)^{1/t} a_\nu \cdot (\nu!)^{-1/t} y^{(\nu)}(x)| \leqq 1.$$

Then from (26) we see that for any solution $y(x)$ of (3) we must have

$$\limsup_{\nu=\infty} |(\nu!)^{-1/t} y^{(\nu)}(x)| \leqq q^{-1}.$$

When $t \leqq 1$ this does not yield any information beyond that implied by the fact that the given solution $y(x)$ is an integral function. But when $t > 1$ it does give such additional information; in fact, it implies that the solution must then be of class $C(1 - t^{-1}, q_1)$ where $0 \leqq q_1 \leqq q^{-1}$. Under the named conditions every solution of (3) must be of the class indicated.

UNIVERSITY OF ILLINOIS.

ON GENERALIZATIONS OF SUM FORMULAS OF THE EULER-MACLAURIN TYPE.

By MARION T. BIRD.

Introduction. We seek solutions $F(x)$ of the difference equation

$$(0.1) \qquad \underset{\omega}{\Delta} F(x) \equiv [F(x+\omega) - F(x)]/\omega = \phi(x)$$

where $\phi(x)$ is a known function. By the method of symbolic operators introduced by Lagrange [1] (see also Pincherle)[2] we see that a formal solution is afforded by the expression

$$F(x+\omega y) = \omega e^{\omega y D}(e^{\omega D} - 1)^{-1}\phi(x).$$

Carmichael [3] has emphasized the fact that by this procedure classes of solutions of the equation (0.1) may be made to depend upon suitable expansions of the expression

$$g(D) = \omega e^{\omega y D}(e^{\omega D} - 1)^{-1}.$$

Such considerations lead to the introduction of the function $A_{\nu,r}(y)$ by means of the Laurent expansion (see Von Koch)[4]

$$(0.2) \quad \omega e^{\omega y z}(e^{\omega z} - 1)^{-1} = \sum_{\nu=-\infty}^{+\infty} \omega^{\nu+1} A_{\nu+1,r}(y) z^{\nu}, \quad 2r\pi < |\omega z| < 2(r+1)\pi,$$
$$(r = 0, 1, 2, \cdots).$$

It is convenient to define $\phi^{(-\nu)}(x)$ by means of the equation

$$\phi^{(-\nu)}(x) = \int_a^x \phi^{(-\nu+1)}(t)\,dt \qquad (\nu = 1, 2, 3, \cdots),$$

where a is in the region of continuity of $\phi(x)$. We may transform $\phi^{(-\nu)}(x)$ into the alternative form

$$\phi^{(-\nu)}(x) = \int_a^x \frac{(x-t)^{\nu-1}}{(\nu-1)!}\phi(t)\,dt, \qquad (\nu = 1, 2, 3, \cdots).$$

[1] J. L. de Lagrange, " Sur une nouvelle espèce de calcul, rélatif à la différentiation et à l'intégration des quantités variables," *Nouveaux Mémoires de L'Académie Royale des Sciences et Belles-Lettres a Berlin* (1772), pp. 185-221.

[2] S. Pincherle, " Funktionaloperationen und -gleichungen," *Encyklopädie der Mathematischen Wissenschaften*, II A 11, pp. 761-817.

[3] R. D. Carmichael, " The present state of the difference calculus and the prospect for the future," *The American Mathematical Monthly*, vol. 31 (1924), pp. 169-183.

[4] Helge Von Koch, " On a class of equations connected with Euler-Maclaurin's sum-formula," *Arkiv för Matematik, Astronomi Och Fysik*, vol. 15 (1921), N:o 26.

With the foregoing definitions in mind one is led to contemplate

$$(0.3)\qquad F(x+\omega y)=\sum_{\nu=-\infty}^{+\infty}\omega^{\nu+1}A_{\nu+1,r}(y)\phi^{(\nu)}(x)\qquad (r=0,1,2,\cdots),$$

as a possible solution of equation (0.1). In the following sections this will be exhibited as an actual solution. However, we shall need to know more properties of the function $A_{\nu,r}(y)$ and the expansions associated with it.

From the generating equation (0.2) it follows that the function $A_{\nu,r}(y)$ satisfies the following relationships:

$$(0.4)\qquad \Delta A_{\nu+1,r}(y)=\left\{\begin{array}{ll}0, & \nu=-1,-2,-3,\cdots\\ 1, & \nu=0.\\ y^\nu/\nu!, & \nu=1,2,3,\cdots\end{array}\right\}\quad (r=0,1,2,\cdots);$$

$$(0.5)\qquad D_yA_{\nu+1,r}(y)=A_{\nu,r}(y)\qquad (\nu=0,\pm 1,\pm 2,\cdots,r=0,1,2,\cdots);$$

$$(0.6)\qquad A_{\nu,r}(1-y)=(-1)^\nu A_{\nu,r}(y)\quad (\nu=0,\pm 1,\pm 2,\cdots,r=0,1,2,\cdots).$$

Furthermore, if we introduce the symbol $[y]$ to represent the largest integer that does not exceed y we have

$$(0.7)\qquad \sum_{s=0}^{p-1}A_{\nu+1,r}(y+sp^{-1})=p^{-\nu}A_{\nu+1,[rp^{-1}]}(py)$$
$$(\nu=0,\pm 1,\pm 2,\cdots,r=0,1,2,\cdots).$$

If C_r denotes a circle with center at the origin and of radius $2\pi r+\delta$, $0<\delta<2\pi$, then it follows from (0.2) that $A_{\nu,r}(y)$ has the alternative definition

$$(0.8)\qquad A_{\nu,r}(y)=(2\pi i)^{-1}\int_{C_r}e^{yt}(e^t-1)^{-1}t^{-\nu}dt$$
$$(\nu=0,\pm 1,\pm 2,\cdots,r=0,1,2,\cdots).$$

This latter equation enables us to relate the function $A_{\nu,r}(y)$ with the Bernoulli-Hurwitz [5] function and enables us to deduce the relation

$$(0.9)\qquad A_{\nu,r}(y)=\left\{\begin{array}{r}\sum_{k=-r}^{+r}(2k\pi i)^{-\nu}e^{2k\pi yi}=2\sum_{k=1}^{r}(2k\pi)^{-\nu}\cos(2k\pi y-\nu\pi/2),\\ (\nu=-1,-2,-3,\cdots),\\ \sum_{k=-r}^{+r}e^{2k\pi yi}=1+2\sum_{k=1}^{r}\cos 2k\pi y=\dfrac{\sin(2r+1)\pi y}{\sin\pi y}\\ (\nu=0),\\ A_{\nu,0}(y)+2\sum_{k=1}^{r}(2k\pi)^{-\nu}\cos(2k\pi y-\nu\pi/2),\\ (\nu=1,2,3,\cdots),\end{array}\right\}$$
$$(r=0,1,2,\cdots).$$

[5] A. Hurwitz, "Sur l'intégrale finie d'une fonction entière," *Acta Mathematica*, vol. 20 (1897), pp. 285-312; vol. 22 (1899), pp. 179-180.

The equation (0.8) also permits us to deduce the following lemma in the manner of Carmichael [6] or Lindelöf: [7]

LEMMA. *For every positive δ less than 2π there exists a constant K_δ, depending on δ alone, such that*

$$|A_{\nu+1,r}(y)| < K_\delta(2r\pi+\delta)^{-\nu}e^{|y|(2r\pi+\delta)}, \quad (\nu = 0, \pm 1, \pm 2, \cdots, r = 0, 1, 2, \cdots),$$

for all finite y.

In the study of the given difference equation for real variables it will be convenient to associate with $A_{\nu,r}(y)$ for real values of y the periodic function $\bar{A}_{\nu,r}(y)$. It is defined by the equation

$$\bar{A}_{\nu,r}(y) = A_{\nu,r}(y - [y]), \qquad (\nu = 1, 2, 3, \cdots, r = 0, 1, 2, \cdots).$$

In order to proceed with rigor we shall have to discuss separately the situations in the real and the complex domains.

I. REAL VARIABLES.

1.1. **The modified Euler-Maclaurin expansion formula.** Let us assume that x, y, z are real variables and that ω is positive. Furthermore, let us use m, $n+1$ to represent positive integers. Let us consider a function $\phi(x)$ which together with its first m derivatives is assumed to be continuous for $b \leqq x \leqq c$, $b + \omega \leqq c$.

Let us contemplate the function

$$P_{\nu,r} = \omega^\nu \int_z^y A_{\nu,r}(y - t + z)\phi^{(\nu)}(x + \omega t)\,dt, \quad (v = -n, -n+1, \cdots, m),$$

where r is a fixed non-negative integer and z, x, y are such that $0 \leqq z \leqq 1$, $b \leqq x \leqq c - \omega$, $b \leqq x + \omega y \leqq c$. Integrating by parts we have at once from (0.5) the relationship

$$P_{\nu,r} = P_{\nu-1,r} + \omega^{\nu-1}\{A_{\nu,r}(z)\phi^{(\nu-1)}(x+\omega y) - A_{\nu,r}(y)\phi^{(\nu-1)}(x+\omega z)\},$$
$$(\nu = -n+1, -n+2, \cdots, m).$$

If we replace ν by $-n+1, -n+2, \cdots, m-1, m$ successively in this last equation and add the resulting equations we have

[6] R. D. Carmichael, "Summation of functions of a complex variable," *Annals of Mathematics* (2), vol. 34 (1933), pp. 349-378.

[7] Ernst Lindelöf, *Le Calcul des Résidus, et ses Applications a la Théorie des Fonctions*, Paris 1905.

$$(1.1)\quad P_{m,r}=P_{-n,r}+\sum_{\nu=-n}^{m-1}\omega^{\nu}\{A_{\nu+1,r}(z)\phi^{(\nu)}(x+\omega y)-A_{\nu+1,r}(y)\phi^{(\nu)}(x+\omega z)\}.$$

If we difference both members of this equation with respect to z and set z equal to zero in the result we have as a consequence of (0.4) an expansion which we incorporate in the following theorem:

THEOREM 1.1. *Let $\phi(x)$ together with its first m, $m\geqq 1$, derivatives be continuous for x in the interval $b\leqq x\leqq c$, $b+\omega\leqq c$. Then $\phi(x)$ has the expansion*

$$(1.2)\qquad \phi(x+\omega y)=\sum_{\nu=-n}^{m-1}\omega^{\nu+1}A_{\nu+1,r}(y)\underset{\omega}{\Delta}\phi^{(\nu)}(x)-R_m+R_{-n}$$

for ω positive, r and n non-negative integers, and for x and y such that $b\leqq x\leqq c-\omega$ and $b\leqq x+\omega y\leqq c$ where

$$R_m=\omega^m\int_0^1\bar{A}_{m,r}(-t)\phi^{(m)}(x+\omega y+\omega t)\,dt$$
$$-\omega^m\int_0^y A_{m,r}(y-t)[\phi^{(m)}(x+\omega+\omega t)-\phi^{(m)}(x+\omega t)]\,dt,$$
$$R_{-n}=\omega^{-n-1}\int_x^{x+\omega}A_{-n,r}(\{x-t+\omega y\}/\omega)\phi^{(-n)}(t)\,dt.$$

We shall designate the expansion (1.2) the *modified Euler-Maclaurin expansion formula* (see Euler,[8] Maclaurin[9]). This expansion is true independently of the value of n and so n may be taken equal to zero without loss of generality.

Derivation of expansion formulas by integration by parts was studied by Darboux[10] but the ideas presented by Nörlund[11] suggested the foregoing derivation.

If z is positive and m is a positive integer all the hypotheses of Theorem 1.1 are met if $\phi(x)=e^{zx}$, $-\infty\leqq x\leqq c-\omega$. In that case we have

$$\phi^{(\nu)}(x)=z^{\nu}e^{xz},\qquad (\nu=0,1,2,\cdots).$$

Furthermore, if a is taken to be negative infinity we have

$$\phi^{(-\mu)}(x)=z^{-\mu}e^{xz},\qquad (\mu=1,2,3,\cdots).$$

[8] Leonhard Euler, "Methodus generalis summandi progressiones," *Commentarii Academiae Scientiarum Imperialis Petropolitanea*, vol. 6 (1732-33), pp. 68-97.

[9] Maclaurin, *A Treatise of Fluxions*, Edinburgh 1742.

[10] G. Darboux, "Sur les développements en série des fonctions d'une seule variable," *Journal de Mathématiques Pures et Appliquées, Paris* (3), vol. 2 (1876), pp. 291-312.

[11] N. E. Nörlund, "Mémoire sur le calcul aux différences finies," *Acta Mathematica*, vol. 44 (1923), pp. 71-211.

Consequently, we have as a special case of Theorem 1.1 the expansion

$$(1.3) \qquad \omega e^{\omega yz}(e^{\omega z}-1)^{-1}=\sum_{\nu=-n}^{m-1}\omega^{\nu+1}A_{\nu+1,r}(y)z^{\nu}+B_m+B_{-n}$$

for ω and z positive, for $r, n, m-1$ non-negative integers and for real finite y where

$$B_m=\omega^{m+1}z^m\int_0^y A_{m,r}(y-t)e^{\omega tz}dt-\omega^{m+1}z^m(e^{\omega z}-1)^{-1}\int_y^{y+1}\bar{A}_{m,r}(y-t)e^{\omega tz}dt,$$

$$B_{-n}=\omega^{-n+1}z^{-n}(e^{\omega z}-1)^{-1}\int_0^1 A_{-n,r}(y-t)e^{\omega tz}dt=\omega\sum_{k=-r}^{+r}\left(\frac{2k\pi i}{\omega z}\right)^n\frac{e^{2k\pi yi}}{\omega z-2k\pi i}.$$

1.2. **The modified Euler-Maclaurin sum formula.** The formula (1.3) might be used in place of (0.2) to suggest a solution of the difference equation (0.1). Using the suggested expansion we are led to contemplate the function

$$\begin{aligned}(1.4) \quad F_r(x+\omega y|\omega)=&\sum_{\nu=-n}^{m-1}\omega^{\nu+1}A_{\nu+1,r}(y)\phi^{(\nu)}(x)\\&+\omega^{m+1}\int_0^y A_{m,r}(y-t)\phi^{(m)}(x+\omega t)dt\\&+\omega^{m+1}\int_0^\infty \bar{A}_{m,r}(-t)\phi^{(m)}(x+\omega y+\omega t)dt\\&+\omega^{-n}\int_a^x A_{-n,r}(\{x-t+\omega y\}/\omega)\phi^{(-n)}(t)dt\end{aligned}$$

as a consequence of the usual interpretation of the symbolic operators involved. As a result of equation (1.1) it can be shown that $F_r(x+\omega y|\omega)$ is actually a function of $x+\omega y$ and, furthermore, that it is independent of the value of n so long as n is a non-negative integer. Moreover, we can prove the theorem:

THEOREM 1.2. *Let $\phi(x)$ together with its first m, $m \geqq 1$, derivatives be continuous for x such that $x \geqq b$ and, moreover, such that for a particular non-negative integral value of r the integral*

$$\int_0^\infty \bar{A}_{m,r}(-t)\phi^{(m)}(x+\omega t)dt$$

converges for x in the interval $b \leqq x \leqq b+\omega$. Then the function $F_r(x+\omega y|\omega)$ exists and affords a solution of the difference equation (0.1) for positive ω and non-negative n and for x and y such that $x \geqq b$ and $x+\omega y \geqq b$.

It is readily seen that the hypotheses imply the existence of $F_r(x+\omega y|\omega)$. That $F_r(x+\omega y|\omega)$ satisfies the given difference equation is a consequence of Theorem 1.1 all of whose hypotheses are satisfied.

We shall designate the solution (1. 4) the *modified Euler-Maclaurin sum formula* (Clearly some distinction must be made between formula (1. 2) and formula (1. 4)). This formula is true independently of the value of n and without loss of generality n may be taken to be zero. For extensive references to the literature of the Euler-Maclaurin formulas one might see Burkhardt,[12] Runge-Willers,[13] Nörlund,[14] and Walther.[15]

If we introduce the function

$$Q_r(x+\omega y)=\omega^{-n}\int_a^x A_{-n,r}(\{x-t+\omega y\}/\omega)\,\phi^{(-n)}(t)\,dt + \sum_{\nu=-n}^{m-1}\omega^{\nu+1}A_{\nu+1,r}(y)\phi^{(\nu)}(x)+\omega^{m+1}\int_0^y A_{m,r}(y-t)\phi^{(m)}(x+\omega t)\,dt$$

it is clear that we may write the function $F_r(x+\omega y|\omega)$ in the form

$$F_r(x+\omega y|\omega)=Q_r(x+\omega y) - \omega\sum_{s=0}^{\infty}\{\phi(x+\omega y+s\omega)-\underset{\omega}{\Delta}\,Q(x+\omega y+s\omega)\} \tag{1.5}$$

so that the modified Euler-Maclaurin sum appears as a particualr modified progressive sum of $\phi(x)$. It is the only solution of (0. 1) having the property

$$\lim_{h\to 0}\{F_r(x+\omega y+h\omega|\omega)-Q_r(x+\omega y+h\omega)\}=0.$$

Let $\phi(x)$ together with its derivatives of all orders be continuous for x such that $x\geqq b$ and, moreover, such that the integrals

$$\int_0^\infty \bar{A}_{\nu,r}(-t)\phi^{(\nu)}(x+\omega t)\,dt, \qquad (\nu=N+1,N+2,\cdots),$$

converge for x in the interval $b\leqq x\leqq b+\omega$ and possess finite limits as ω approaches zero. Then, considered as a function of ω, $F_r(x+\omega y|\omega)$ is such that

$$F_r(x+\omega y|\omega)\sim\omega^{-n}\int_a^x A_{-n,r}(\{x-t+\omega y\}/\omega)\phi^{(-n)}(t)\,dt + \sum_{\nu=-n}^{m-1}\omega^{\nu+1}A_{\nu+1,r}(y)\phi^{(\nu)}(x), \qquad \omega>0.$$

---

[12] H. Burkhardt, "Restglied der Euler-Maclaurinschen Summenformel," *Encyklopädie der Mathematischen Wissenschaften*, II A 12, Nr. 105, pp. 1324-1337.

[13] C. Runge-Fr. A. Willers, "Die Eulersche Formel," *Encyklopädie der Mathematischen Wissenschaften*, II C 2, Nr. 9, pp: 91-96.

[14] N. E. Nörlund, "Einfache Summen," *Encyklopädie der Mathematischen Wissenschaften*, II C 7, Nr. 10, pp. 711-716.

[15] Alwin Walther, "Numerische Integration," *Pascals Repertorium der Höhern Mathematik*, 2 Auflage 1929, I 3 XXIII, 4, pp. 1200-1210.

As a consequence of the property (0. 7) of the Bernoulli-Hurwitz function the modified Euler-Maclaurin sum enjoys the relation

$$\sum_{s=0}^{p-1} F_r(x+\omega s p^{-1}|\omega) = pF_{[rp^{-1}]}(x|\omega p^{-1}) \tag{1.6}$$

whenever the left member exists.

Let $\phi(x)$ together with its first m, $m \geqq 1$, derivatives be continuous for x such that $x \geqq b$ and, moreover, such that for a particular non-negative value of r the integral

$$\int_0^\infty \bar{A}_{m,r}(-t)\phi^{(m)}(x+\omega t)\,dt$$

converges uniformly for x such that $b \leqq x \leqq b+\omega$. Then we have

$$\omega^{-1}\int_x^{x+\omega} F_r(t|\omega)\,dt = \int_a^x \phi(t)\,dt. \tag{1.7}$$

Let $\phi(x)$ together with its first m, $m \geqq 1$, derivatives be continuous for x such that $x \geqq b$ and, moreover, such that the series

$$\sum_{s=0}^{\infty} \phi^{(m)}(x+s\omega)$$

converges uniformly in the interval $b \leqq x \leqq b+\omega$. Then, if $F_r{}^{\mu}(x|\omega)$ denotes the modified Euler-Maclaurin sum of $\phi^{(\mu)}(x)$, the first m derivatives of the modified Euler-Maclaurin sum of $\phi(x)$ exist and are such that

$$F_r{}^{(\mu)}(x|\omega) = F_r{}^{\mu}(x|\omega) + \sum_{\nu=0}^{\mu-1} \omega^{-\nu} A_{-\nu,r}(\{x+\omega y - a\}/\omega)\phi^{(\mu-\nu-1)}(a), \tag{1.8}$$
$$(\mu = 0, 1, \cdots, m).$$

Under the conditions named in the foregoing paragraph together with the condition that $\phi(x)$ have a zero of order m at some point a, $a \geqq b$, the modified Euler-Maclaurin sum of $\phi^{(\mu)}(x)$ equals the μ-th derivative of the modified Euler-Maclaurin sum of $\phi(x)$.

The proofs of the theorems listed above follow from Theorem 1. 2 and the properties of the Bernoulli-Hurwitz function by allowing y to be the variable of integration, summation, and differentiation.

As an example of the modified Euler-Maclaurin sum let us take $\phi(x)$ equal to $x^\mu/\mu!$, $\mu \geqq 0$. It is clear that all the conditions of Theorem 1. 2 are satisfied if m is taken to be $\mu+1$. Then if a is taken to be zero we have

$$F_r(x|\omega) = \omega^{\mu+1}A_{\mu+1,r}(x/\omega)$$

and the Bernoulli-Hurwitz function itself is seen to be the sum of x^μ afforded by the modified Euler-Maclaurin sum formula. This leads us to expect the Euler-Maclaurin sum is intimately related to a modification of Nörlund's [11] principal sum. We propose to indicate this relationship more precisely in the next section.

1. 3. **Modified principal solutions.** Under suitable hypotheses $\phi(x)$ has the expansion

$$\phi(x+s\omega)=\omega^{-1}\int_{s\omega}^{s\omega+\omega}\phi(t)\,dt+2\omega^{-1}\sum_{k=1}^{\infty}\int_{s\omega}^{s\omega+\omega}\cos[2k\pi(x-t)/\omega]\phi(t)\,dt$$

and consequently $\phi(x+s\omega)$ is approximated by the quantity

$$\omega^{-1}\int_{s\omega}^{s\omega+\omega}\phi(t)\,dt+2\omega^{-1}\sum_{k=1}^{r}\int_{s\omega}^{s\omega+\omega}\cos[2k\pi(x-t)/\omega]\phi(t)\,dt$$
$$=\omega^{-1}\int_{s\omega}^{s\omega+\omega}A_{0,r}(\{x-t\}/\omega)\phi(t)\,dt.$$

Now a solution of (0. 1) is afforded by the series

$$-\omega\sum_{s=0}^{\infty}\phi(x+s\omega) \tag{1.9}$$

if it converges; if it diverges one might naturally contemplate the difference

$$\int_a^\infty A_{0,r}(\{x-t\}/\omega)\phi(t)\,dt-\omega\sum_{s=0}^{\infty}\phi(x+s\omega)$$

in order to find a valid solution of (0. 1) (see Kronecker).[16] Nörlund [11] has contemplated the difference for the special case where r is zero.

If we follow the procedure which Nörlund has given for the special case we would define the modified principal solution $F_r(x|\omega)$ to be the limit of $F_r(x|\omega;\eta)$ as η approaches zero where

$$F_r(x|\omega;\eta)=\int_a^\infty A_{0,r}(\{x-t\}/\omega)e^{-\eta t}\phi(t)\,dt-\omega\sum_{s=0}^{\infty}\phi(x+s\omega)e^{-\eta(x+s\omega)},\quad \eta>0.$$

The relations (1. 6) and (1. 7) which held for the modified Euler-Maclaurin sum are seen to hold for the modified principal solution as a consequence of this definition.

Furthermore, the procedure of Nörlund leads to the theorem:

---

[16] Leopold Kronecker, "Ueber eine bei Anwendung der partiellen Integration nützliche Formel," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* (1885), pp. 841-862; *Leopold Kronecker's Werke,* vol. 5 (1930), pp. 267-294.

THEOREM 1.3. *Let $\phi(x)$ together with its first m, $m \geqq 1$, derivatives be continuous for $x \geqq b$ and such that for a fixed value of r the integral*

$$\int_0^\infty \bar{A}_{m,r}(-t)\phi^{(m)}(x+\omega t)\,dt$$

converges uniformly for x such that $b \leqq x \leqq b+\omega$. Then the modified principal solution exists as a continuous function of x and is afforded by the modified Euler-Maclaurin sum formula

$$F_r(x+\omega y|\omega) = \int_a^x A_{0,r}(\{x-t+\omega y\}/\omega)\phi(t)\,dt + \sum_{\nu=1}^{m} A_{\nu,r}(y)\omega^\nu\phi^{(\nu-1)}(x) + \omega^{m+1}\int_0^\infty \bar{A}_{m,r}(y-t)\phi^{(m)}(x+\omega t)\,dt$$

for $x \geqq b$, $0 \leqq y \leqq 1$ provided $a \geqq b$.

Solutions valid for $x \leqq b$ would be obtained by proceeding from the series

$$\omega \sum_{s=1}^{\infty} \phi(x - s\omega)$$

rather than from the series (1.9).

II. COMPLEX VARIABLES.

We shall confine attention to integral functions. A classification of integral functions developed in an unpublished paper by Carmichael, Martin, and Bird is used.

2.1. **A classification of integral functions.** If the sequence of positive numbers $t_0, t_1, t_2, \cdots$ has, in addition to the property that

$$\lim_{n\to\infty} (t_n)^{1/n} = \infty, \tag{2.1}$$

the property that for every analytic function $F(x)$ we have

$$\limsup_{n\to\infty} |\, t_n F^{(n)}(x_0)/n!\,|^{1/n} = \limsup_{n\to\infty} |\, t_n F^{(n)}(x_1)/n!\,|^{1/n}$$

for any two regular points x_0 and x_1 of $F(x)$, then we shall say that $\{t_n\}$ is an I-*sequence*. If $\{t_n\}$ is an I-sequence and $F(x)$ is an analytic function such that at the regular point x_0 we have

$$\limsup_{n\to\infty} |\, t_n F^{(n)}(x_0)/n!\,|^{1/n} = q, \qquad 0 \leqq q < \infty,$$

then $F(x)$ is an integral function and we shall say it is a function of *sort* $\{t_n\}$, *type* q.

It is convenient to introduce the particular function $E(x)$ of sort $\{t_n\}$, type 1, defined by the expansion

$$E(x) = \sum_{\nu=0}^{\infty} x^\nu / t_\nu.$$

It is also useful to introduce the sequences $\{\lambda_n\}$ and $\{\rho_n\}$ such that

$$\lambda_0 = t_0,\ \lambda_1 = t_1,\ \lambda_n = t_n/t_{n-1}, \qquad (n = 2, 3, 4, \cdots),$$
$$\rho_0 = t_0,\ \rho_n = (t_n)^{1/n}, \qquad (n = 1, 2, 3, \cdots).$$

We have the following theorems:

THEOREM 2.1. *A necessary and sufficient condition that a sequence $\{t_n\}$ shall be an* I*-sequence is that it shall satisfy condition* (2.1) *and the condition*

$$\lim_{n\to\infty} (t_n E^{(n)}(a)/n!)^{1/n} = 1.$$

COROLLARY. *A necessary condition that $\{t_n\}$ be an* I*-sequence is that*

$$\lim_{n\to\infty} \rho_n = \infty.$$

A sufficient condition that $\{t_n\}$ be an I*-sequence is that*

$$\lim_{n\to\infty} \lambda_n = \infty.$$

THEOREM 2.2. *Let $F(x)$ be any integral function. If $F(x)$ is a polynomial then $F(x)$ is of sort $\{t_n\}$, type* 0, *for every* I*-sequence $\{t_n\}$. If $F(x)$ is not a polynomial then there is associated with $F(x)$ an* I*-sequence $\{t_n\}$ such that $F(x)$ is of sort $\{t_n\}$, type* 1.

COROLLARY. *In order to classify all integral functions it is sufficient to consider only the* I*-sequences $\{t_n\}$ for which ρ_n is monotonic.*

THEOREM 2.3. *Let $\{t_n\}$ be an* I*-sequence such that ρ_n is monotonic. Then if $F(x)$ is of sort $\{t_n\}$, type q, we have for every positive ϵ an M_ϵ such that*

$$| t_n F^{(n)}(x)/n! | < M_\epsilon (q+\epsilon)^n (1 + 1/\delta)^n E([q+\epsilon][1+\delta] \, |x|)$$

for every finite x and positive δ.

THEOREM 2.4. *Let $F(x)$ be a function of sort $\{t_n\}$, type q. It is necessary and sufficient in order that the derivative of $F(x)$ be a function of sort $\{t_n\}$, type q, that*

$$\lim_{n\to\infty} (\lambda_n)^{1/n} = 1.$$

2.2. **The generalized Euler-Maclaurin expansion formula.** We consider x, y, z and ω as complex variables. We shall assume the line integrals are taken along straight line paths. Let us assume that m and μ are non-negative integers such that μ equals or exceeds m.

As in Section 1.1 let us consider the function

$$P_{\nu,r} = \omega^\nu \int_z^y A_{\nu,r}(y - t + z)\phi^{(\nu)}(x + \omega t)\,dt, \quad (\nu = 0, 1, 2, \cdots, r = 0, 1, 2, \cdots).$$

If we integrate by parts and then replace ν by $m + 1, m + 2, \cdots, \mu + 1$ in the result we have upon addition

$$P_{\mu+1,r} = P_{m,r} + \sum_{\nu=m}^{\mu} \omega^\nu \{A_{\nu+1,r}(z)\phi^{(\nu)}(x + \omega y) - A_{\nu+1,r}(y)\phi^{(\nu)}(x + \omega z)\}, \quad (r = 0, 1, 2, \cdots).$$

Let us difference both members of this equation with respect to z and then set z equal to zero. For abbreviation we use

$$R_{\nu,r} = [\Delta P_{\nu,r}]_{z=0}, \qquad (\nu = 0, 1, 2, \cdots, r = 0, 1, 2, \cdots).$$

Then we have the formula

$$\sum_{\nu=m}^{\mu} \omega^{\nu+1} A_{\nu+1,r}(y) \underset{\omega}{\Delta} \phi^{(\nu)}(x) + R_{\mu+1,r} - R_{m,r} = \left\{ \begin{array}{ll} 0, & m > 0, \\ \phi(x + \omega y), & m = 0 \end{array} \right\} \quad (r = 0, 1, 2, \cdots).$$

Let the integer r depend upon ν for its value and let us designate this relationship by the notation r_ν, $\nu = 0, 1, 2, \cdots$. If we have

$$r_\nu = r_m, \qquad (\nu = m, m + 1, \cdots, \mu),$$

the preceding formula assumes the form

$$\sum_{\nu=m}^{\mu} \omega^{\nu+1} A_{\nu+1,r_\nu}(y) \underset{\omega}{\Delta} \phi^{(\nu)}(x) + R_{\mu+1,r_\mu} - R_{m,r_m} = \left\{ \begin{array}{ll} 0, & m > 0, \\ \phi(x + \omega y), & m = 0. \end{array} \right.$$

Let us arrange successive values of r_ν which are equal in groups. Let us designate the number in successive groups by ν_1, $\nu_2 - \nu_1$, $\nu_3 - \nu_2, \cdots$. It is convenient to adopt the convention that

$$\sum_{k=r+1}^{\rho} f(k) = \sum_{k=0}^{\rho} f(k) - \sum_{k=0}^{r} f(k), \qquad r \geqq \rho \geqq 0.$$

Now let us take m equal successively to $0, \nu_1, \cdots, \nu_j$ where

$$\nu_j \leqq n < \nu_{j+1}$$

and let us take μ equal successively to $\nu_1 - 1, \nu_2 - 1, \cdots, n$. If we add the resulting expansions we obtain as a consequence of the properties of the Bernoulli-Hurwitz functions the expansion

$$(2.2) \qquad \phi(x + \omega y) = I_{n+1,r_n} + \sum_{\nu=0}^{n} \omega^{\nu+1} A_{\nu+1,r_\nu}(y) \underset{\omega}{\Delta} \phi^{(\nu)}(x) + R_{n+1,r_n}$$

where

$$(2.3) \qquad R_{n+1,r_n} = \omega^{n+1}\{\int_1^y A_{n+1,r_n}(y - t + 1)\phi^{(n+1)}(x + \omega t)\,dt - \int_0^y A_{n+1,r_n}(y - t)\phi^{(n+1)}(x + \omega t)\,dt\},$$

$$(2.4) \qquad I_{n+1,r_n} = \int_0^1 \{\phi(x + \omega t) + 2 \sum_{\nu} \sum_{k=r_{\nu-1}+1}^{r_\nu} \omega^\nu (2k\pi)^{-\nu} \cos[2k\pi(y - t) - \nu\pi/2]\phi^{(\nu)}(x + \omega t)\}dt.$$

In this latter integral ν runs over the values $0, \nu_1, \nu_2, \cdots, \nu_j$ with r_{-1} equal to zero.

We shall refer to (2.2) as the *generalized Euler-Maclaurin expansion.*

Let us assume that $\phi(x)$ is of sort $\{t_n\}$, type q, where $\{t_n\}$ is an I-sequence such that ρ_n is monotonic. Let us assume that ω is held fixed while x and y are confined to the finite regions X and Y of the complex plane. Then the lengths of the paths of the integrals in (2.2) as well as the absolute values of each of the quantities $y - t + 1$, $y - t$, $x + \omega t$ are dominated by the constant a. Furthermore, as a consequence of Theorem 2.3 and the Lemma in the Introduction we have

$$\limsup_{n\to\infty} (R_{n+1,r_n})^{1/n} < 1$$

if r_n is suitably restricted. Indeed it is sufficient that we have

$$q\,|\,\omega\,|\,(1 + \gamma)/(2\pi e\rho_{n+1}) \leqq (r_n + 1)/(n + 1) \leqq [\log \rho_{n+1}]^\alpha$$

where α is any positive constant less than 1 and γ is any positive constant. The proof is facilitated if a division of cases is made at

$$r_n + 1 = (n + 1)\xi$$

where ξ is any positive constant which satisfies the inequality

$$2\pi a\xi - \log(1 + \gamma) < 0.$$

If we restrict the sequence $\{t_n\}$ to be such that

$$\lim_{n\to\infty} (\lambda_n)^{1/n} = 1,$$

under the conditions named above we also have

$$\limsup_{n\to\infty} | \omega^{n+1} A_{n+1,r_n}(y) \underset{\omega}{\Delta} \phi^{(n)}(x) |^{1/(n+1)} < 1.$$

Consequently, we have the theorem:

THEOREM 2.5. *Let $\phi(x)$ be a function of sort $\{t_n\}$, type q, where $\{t_n\}$ is an I-sequence such that ρ_n is monotonic and*

$$\lim_{n\to\infty} (\lambda_n)^{1/n} = 1.$$

Then for x, y confined to finite regions X, Y and for a fixed value of ω, $\phi(x)$ has the expansion

$$\phi(x+\omega y) = \lim_{n\to\infty} I_{n+1,r_n} + \sum_{\nu=0}^{\infty} \omega^{\nu+1} A_{\nu+1,r_\nu}(y) \underset{\omega}{\Delta} \phi^{(\nu)}(x) \tag{2.5}$$

where for all n greater than some fixed value we have

$$q \mid \omega \mid (1+\gamma)/(2\pi e \rho_{n+1}) \leqq (r_n+1)/(n+1) \leqq [\log \rho_{n+1}]^{\alpha},$$
$$0 < \alpha < 1, \qquad 0 < \gamma,$$

and where the integral I_{n+1,r_n} is defined in the preceding remarks. This expansion converges uniformly in the regions indicated. Furthermore, this expansion is unique for any particular sequence $\{r_n\}$.

In particular, one may choose r_n to be n. This is the choice Hurwitz [5] made in his study of the equation (0.1) for integral functions.

For functions of exponential type q (see Carmichael),[17] i. e., $\lambda_n = n$, it is sufficient according to Theorem 2.5 to define r_n by the inequalities

$$(q \mid \omega \mid /2\pi) - 1 < r_n \leqq q \mid \omega \mid /2\pi.$$

In such a case Theorem 2.5 recovers the theorem of Carmichael [6] for expansions of functions of exponential type q in series of Bernoulli-Hurwitz functions. For the case where $\lambda_n = n^{1/t}$, $1 \leqq t < \infty$, it is sufficient that

$$(q \mid \omega \mid /2\pi) n^{1-1/t} < r_n \leqq (q \mid \omega \mid /2\pi) n^{1-1/t}, \qquad (n = 1, 2, 3, \cdots).$$

When q is different from zero this gives expansions of functions of order t, normal type q (see Bieberbach).[18]

---

[17] R. D. Carmichael, "Functions of exponential type," *Bulletin of the American Mathematical Society*, vol. 41 (1934), pp. 241-261.

[18] Ludwig Bieberbach, "Grundbegriffe," *Encyklopädie der Mathematischen Wissenschaften*, II C 4, Nr. 29, pp. 429-431.

2.3. **The generalized Euler-Maclaurin sum formula.** In this section we propose to modify the results of Guichard,[19] Appell,[20] Hurwitz,[5] Weber,[21] Barnes,[22] Carmichael [23] so that instead of finding an integral solution of the equation (0.1) when the known function is integral we shall find an integral solution of the same character as the known function.

We state the result in two theorems:

THEOREM 2.6. *Let $\phi(x)$ be an integral function of sort $\{t_n\}$, type q, $0 < q < \infty$, where $\{t_n\}$ is an* I*-sequence such that*

$$\lim_{n\to\infty} \lambda_n = \infty, \qquad \limsup_{n\to\infty} (\lambda_{n+1}/\lambda_n)^n = \theta e < e.$$

Then there exists a solution of the difference equation (0.1) *which is of sort $\{t_n\}$, type q, and indeed one such solution is afforded by the function*

$$F(x|\omega) = \sum_{\nu=0}^{\infty} \phi^{(\nu)}(0) A_{\nu+1,r_\nu}(x/\omega)\omega^{\nu+1}, \tag{2.6}$$

where r_n is the integer defined by the relation

$$r_n = [q \mid \omega \mid n/2\pi\lambda_n], \qquad (n = 0, 1, 2, \cdots).$$

THEOREM 2.7. *Let $\phi(x)$ be an integral function of sort $\{t_n\}$, type q, $0 < q < \infty$, where $\{t_n\}$ is an* I*-sequence such that*

$$\lim_{n\to\infty} n/\lambda_n = \tau < \infty.$$

Then there exists a solution of the difference equation (0.1) *which is of sort $\{t_n\}$, type q, and indeed one such solution is afforded by*

$$F(x|\omega) = \sum_{\nu=0}^{\infty} \phi^{(\nu)}(0) A_{\nu+1,r}(x/\omega)\omega^{\nu+1} \tag{2.7}$$

[19] G. Guichard, "Sur la résolution de l'équation aux différences finies $G(x+1) - G(x) = H(x)$," *Annales Scientifiques de l'École Normale Supérieure* (3), vol. 4 (1887), pp. 361-380.

[20] P. Appell, "Sur les fonctions périodiques de deux variables," *Journal de Mathématiques Pures et Appliquées, Paris* (4), vol. 7 (1891), pp. 157-219.

[21] Heinrich Weber, "Über Abel's Summation endlicher Differenzenreihen," *Acta Mathematica*, vol. 27 (1903), pp. 225-233.

[22] E. W. Barnes, "The linear difference equation of the first order," *Proceedings of the London Mathematical Society* (2), vol. 2 (1904), pp. 439-469.

[23] R. D. Carmichael, "On the theory of linear difference equations," *American Journal of Mathematics*, vol. 35 (1913), pp. 163-182.

where r is the integer defined by the relation

$$r = [q \mid \omega \mid \tau/2\pi].$$

It is clear that Theorem 2. 6 for the special case $\lambda_n = n^{1/t}$, $1 \leqq t < \infty$ refines the result which Whittaker [24] has given for functions of order t.

Obviously, if the series (2. 6) converges it defines a solution of (0. 1). It follows from the hypotheses of Theorem 2. 6 that we have

$$\lim_{n\to\infty} n/\lambda_n = \infty.$$

Consequently, r_n becomes infinite with n (although not necessarily monotonically) and for any positive numbers δ, δ' we have the inequalities

$$\begin{aligned} 2\pi r_n + \delta &< (q \mid \omega \mid n/\lambda_n) + 2\delta < (q + 2\delta) \mid \omega \mid n/\lambda_n \\ 2\pi(r_n + 1) - \delta' &> (q \mid \omega \mid n/\lambda_n) - \delta' > (q - \delta') \mid \omega \mid n/\lambda_n \end{aligned}$$

for all n after some fixed value. This together with the Lemma of the Introduction and Theorem 2. 3 enables us to prove that the series (2. 6) converges absolutely and uniformly for x confined to a finite region X.

Consideration of the successive derivatives of the integral function defined by the series (2. 6) proves that the function is of sort $\{t_n\}$, type not exceeding q. We are led to a contradiction of our hypotheses if $F(x|\omega)$ is of sort $\{t_n\}$, type less than q. This establishes Theorem 2. 6.

The proof of Theorem 2. 7 follows the same method.

COROLLARY. *The solutions $F(x|\omega)$ defined in Theorem* 2. 6 *and Theorem* 2. 7 *are such that we have*

$$F(x-\omega|-\omega) = F(x|\omega).$$

This is an immediate consequence of the relation (0. 6).

2. 4. **Modified principal solutions.** In a manner analogous to that found in Section 1. 3 we may modify Nörlund's definition of principal solution in the complex domain. With this modification many of Nörlund's results may be extended. In particular, we have the theorem:

THEOREM 2. 8. *Let $\phi(x)$ be an integral function of exponential type q, $0 \leqq q < \infty$. Then the modified principal solution $F_r(x|\omega)$ exists as an analytic function of x for all finite x and of ω for all ω in the interior of the circle*

[24] J. M. Whittaker, "On the asymptotic periods of integral functions," *Proceedings of the Edinburgh Mathematical Society* (2), vol. 3 (1932-33), pp. 241-258.

with center at the origin and of radius $2(r+1)\pi/q$ *with a neighborhood about the origin deleted if* $r>0$ *(If* $q=0$*, the region of validity is the* ω*-plane with a neighborhood about the origin deleted). Furthermore, if* ω *is confined to the annular ring*

$$2r\pi \leqq |\omega| q < 2(r+1)\pi,$$

the modified principal solution is of exponential type q.

COROLLARY. *Under the conditions named in the preceding theorem the modified principal solution may be exhibited in the forms:*

$$\begin{aligned}
F_r(x|\omega) &= \int_a^{x-\alpha\omega} A_{0,r}(\{x-t\}/\omega)\phi(t)\,dt \\
&\quad + \omega\int_{-\alpha}^{-\alpha+\infty i} \phi(x+\omega t)e^{2r\pi ti}(1-e^{-2\pi ti})^{-1}\,dt \\
&\quad + \omega\int_{-\alpha}^{-\alpha-\infty i} \phi(x+\omega t)e^{-2r\pi ti}(1-e^{2\pi ti})^{-1}\,dt, \\
F_r(x|\omega) &= (2r+1)(2\pi i)^{-1}\int_{C_2}\int_a^{x+\omega t} A_{0,r}(\{x-t\}/\omega)\phi(u)\,du \\
&\quad \times \{\pi\csc(2r+1)\pi t\}^2\,dt,
\end{aligned}$$

$$\begin{aligned}
F_r(x|\omega) &= \int_a^{x} A_{0,r}(\{x-t\}/\omega)\phi(t)\,dt - \omega\phi(x)2^{-1} \\
&\quad + i\omega\int_0^{\infty}\{\phi(x+i\omega t)-\phi(x-i\omega t)\}e^{-2r\pi t}(1-e^{2\pi t})^{-1}\,dt,
\end{aligned} \tag{2.8}$$

$$\begin{aligned}
F_r(x|\omega) &= \int_a^{x} A_{0,r}(\{x-t\}/\omega)\phi(t)\,dt \\
&\quad - \omega\phi(x)2^{-1} + \sum_{\nu=1}^{\infty} A_{\nu+1,r}(0)\omega^{\nu+1}\phi^{(\nu)}(x),
\end{aligned}$$

$$F_r(x|\omega) = \sum_{\nu=1}^{\infty}\omega^{\nu}\phi^{(\nu-1)}(0)A_{\nu+1,r}(x/\omega), \tag{2.9}$$

where a *is an arbitrary constant,* α *is a positive constant less than one, and* C_2 *is the contour through the point* $-\alpha$ *along a line parallel to the axis of imaginaries and traversing it in the negative sense. In the last formula we have taken the arbitrary constant equal to zero.*

We might refer to (2.8) as the *modified Abel sum formula* (see Abel,[25] also Plana).[26]

[25] Niels Henrik Abel, "Solution de quelques problèmes à l'aide d'intégrales définies," *Magazin for Naturvidenskaberne, Christiana* (1), vol. 2 (1823), pp. 55-63; *Oeuvres Complètes de Niels Henrik Abel,* Ed. 2, vol. 1, Christiana (1881), pp. 11-27.

[26] Plana, "Note sur une nouvelle expression analytique des nombres Bernoulliens, propre à exprimer en termes finis la formule générale pour la sommation des suites," *Memorie della Reale Accademia delle Scienze di Torino,* vol. 25 (1820), pp. 403-418.

The formula (2. 9) was exhibited as a solution of the difference equation (0. 1) in Theorem 2. 7. In this connection one is led to contemplate as a possible solution of equation (0. 1) the function defined by the series

$$F_r(x+\omega y|\omega)=\int_a^x A_{0,r}(\{x-t+\omega y\}/\omega)\phi(t)\,dt+\sum_{\nu=1}^{\infty}A_{\nu,r}(y)\omega^{\nu}\phi^{(\nu-1)}(x),\qquad (r=0,1,2,\cdots).$$

If $\phi(x)$ is of exponential type q this function exists as an analytic function of x, y, ω. If we restrict ω to the annular ring

$$2r\pi\leqq|\omega|\,q<2(r+1)\pi$$

it is seen that $F_r(x+\omega y|\omega)$ affords a solution of exponential type q considered as a function of x or of y. This can be established independently of the definition of principal solution.

It is well to note that the modified principal solution extends the range of validity of solutions in the ω-plane. This is rather nicely exhibited by consideration of the particular function $\phi(x)=e^{tx}$.

The extension of our results to functions of two or more variables in the way in which Nörlund,[27] Baten,[28] and others have led is direct but the results become very involved.

SOUTHWESTERN COLLEGE,
WINFIELD, KANSAS.

[27] N. E. Nörlund, "Mémoire sur les polynomes de Bernoulli," *Acta Mathematica*, vol. 43 (1922), pp. 121-196.

[28] W. D. Baten, "A remainder for the Euler-Maclaurin summation formula in two independent variables," *American Journal of Mathematics*, vol. 54 (1932), pp. 265-275.

ON A FUNDAMENTAL THEOREM IN MATRIC THEORY.[1]

By C. C. MacDuffee.

1. The following theorem was first proved by Muth[2] when A is a bilinear form in $2n$ contragradient variables. It was given for a general matrix by Kreis,[3] and by a number of later writers.[4] It is believed, however, that the present proof is the shortest.

Let A be a matrix of order n with elements in a field $\mathfrak{F}^*$ having the single elementary divisor $(\lambda - a)^n$ where a is in $\mathfrak{F}^*$. Let ϕ be any polynomial with coefficients in $\mathfrak{F}^*$.

Theorem 1. *If the first number of the sequence*

$$\phi'(a), \phi''(a), \cdots, \phi^{(n-1)}(a), 1$$

which is not 0 *is the i-th, define q and r by the relations*

$$n = qi + r \qquad 0 \leqq r < i.$$

Then $\phi(A)$ *has the elementary divisors* $[\lambda - \phi(a)]^{q+1}$ *taken r times and* $[\lambda - \phi(a)]^q$ *taken* $i - r$ *times.*

Let A be taken in the Jordan normal form. Then

$$\phi(A) = \left\| \begin{matrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ 0 & a_0 & a_1 & \cdots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & a_0 \end{matrix} \right\| \tag{1}$$

where $a_0 = \phi(a)$, $a_i = (1/i!)\phi^{(i)}(a)$. Using the notation of Turnbull and Aitken,[5] write

$$\phi(A) = a_0 I + a_1 U + a_2 U^2 + \cdots + a_{n-1} U^{n-1}$$

where

$$U = (\delta_{r+1,s}), \qquad U^i U^j = U^{i+j}, \qquad U^n = 0.$$

---

[1] Presented to the American Mathematical Society, April 10, 1936.

[2] P. Muth, *Journal für reine und angewandte Mathematik*, vol. 125 (1903), p. 291.

[3] H. Kreis, *Contribution a la théorié des systémes linéaires*, Zurich, 1906, p. 47.

[4] A number of other references are given by N. H. McCoy, *American Journal of Mathematics*, vol. 57 (1935), pp. 491-502.

[5] Turnbull and Aitken, *Canonical Matrices*, Blackie and Son, 1932, p. 62.

If $a_1 = a_2 = \cdots = a_{i-1} = 0$, $a_i \neq 0$, then

$$\phi(A) - a_0 I = a_i U^i + \cdots, \qquad [\phi(A) - a_0 I]^k = a_i^k U^{ik} + \cdots.$$

The nullity (order minus rank) of U^h is h, so the nullities of the powers of $\phi(A) - a_0 I$ are

$$i, 2i, 3i, \cdots, qi, n,$$

and the Weyr characteristic[6] of $\phi(A)$ relative to the root $a_0 = \phi(a)$ is $(i, i, \cdots, i, r)$. The Segre characteristic of $\phi(A)$ is the conjugate partition of n, namely $(q+1, \cdots, q+1, q, \cdots, q)$ where the number of $q+1$'s is r. This proves the theorem.

2. McCoy[7] has recently extended this theorem to matrices having elements in a general field. The following proof, using results from algebraic number theory, is more direct.

Suppose that $\mathfrak{F}$ is any field, and that the matrix A of order n has elements in $\mathfrak{F}$ and that the polynomial ϕ has coefficients in $\mathfrak{F}$. In the polynomial ring $\mathfrak{F}[\lambda]$, let A have the single elementary divisor $[p(\lambda)]^k$ where $p(\lambda)$ is irreducible of degree h, $n = hk$.

Let $\Phi(\lambda) = 0$ be the equation of degree h obtained by applying ϕ as a Tschirnhausen transformation to the roots of $p(\lambda) = 0$. That is, the roots of $\Phi = 0$ are the functions $\phi(\lambda)$ of the roots of $p(\lambda) = 0$. Then

$$\Phi(\lambda) = [\psi(\lambda)]^m, \tag{2}$$

where $\psi(\lambda)$ has coefficients in $\mathfrak{F}$, is irreducible in $\mathfrak{F}$, and $h = mt$.

THEOREM 2. *If the first function of the sequence*

$$\phi'(\lambda), \phi''(\lambda), \cdots, \phi^{(k-1)}(\lambda), 1$$

which is not divisible by $p(\lambda)$ is the i-th, define q and r by the relations

$$k = qi + r \qquad 0 \leqq r < i.$$

Then $\phi(A)$ has the elementary divisors $[\psi(\lambda)]^{q+1}$ taken mr times and $[\psi(\lambda)]^q$ taken $m(i-r)$ times, where $\psi(\lambda)$ is defined by (2).

Consider $\mathfrak{F}$ extended to a field $\mathfrak{F}^*$ in which

$$p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_h).$$

---

[6] Turnbull and Aitken, *loc. cit.*, p. 80.

[7] N. H. McCoy, *loc. cit.* McCoy's treatment is more explicit in actually exhibiting the transforming matrix.

Then A has the elementary divisors $(\lambda-\lambda_j)^k$. Since $p(\lambda)$ is irreducible, $\phi^{(g)}(\lambda_j)=0$ if and only if $p(\lambda)$ divides $\phi^{(g)}(\lambda)$. Thus for each elementary divisor $(\lambda-\lambda_j)^k$, the same i, q and r are obtained. Then by Theorem 1 the elementary divisors of $\phi(A)$ relative to $\mathfrak{F}^*$ are $[\lambda-\phi(\lambda_j)]^{q+1}$ repeated r times and $[\lambda-\phi(\lambda_j)]^q$ repeated $i-r$ times, $(j=1,2,\cdots,h)$.

It may not be true that the $\phi(\lambda_i)$ are distinct, but if they are not, they fall into t sets of $m=h/t$ equal values each, and the product

$$\psi(\lambda)=\Pi[\lambda-\phi(\lambda_j)],$$

where $\phi(\lambda_j)$ ranges over t distinct values, has coefficients in $\mathfrak{F}$ and is irreducible in $\mathfrak{F}$.[8]

The invariant factors of $\phi(A)$ are the same for $\mathfrak{F}$ as for $\mathfrak{F}^*$, and by the usual rule for forming them from the elementary divisors there are mr of them of the form $[\psi(\lambda)]^{q+1}$ and $m(i-r)$ of them of the form $[\psi(\lambda)]^q$. Since $\psi(\lambda)$ is irreducible in $\mathfrak{F}$, these are the elementary divisors of $\phi(A)$ relative to $\mathfrak{F}$.

The function $\psi(\lambda)$ is readily calculated as follows. Let B be the companion matrix of $p(\lambda)=0$. Then $\Phi(\lambda)=|\phi(B)-\lambda I|$.[9] Then $\psi(\lambda)$ is $\Phi(\lambda)$ divided by the g. c. d. of $\Phi(\lambda)$ and $\Phi'(\lambda)$.

THE UNIVERSITY OF WISCONSIN.

[8] See any standard work on algebraic number theory, as E. Landau, *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale,* Teubner 1918, p. 11, Theorem 33. Or L. W. Reid, *Elements of the Theory of Algebraic Numbers,* Macmillan 1910, p. 273.

[9] C. C. MacDuffee, "The theory of matrices," *Ergeb. Math.*, II, vol. 5, Springer 1933, p. 24.

A GENERALIZED LAMBERT SERIES.[1]

By Mary Cleophas Garvin.

I. *Introduction.* In 1771 Lambert[2] introduced the series

$$\sum_{n=1}^{\infty} \frac{z^n}{1-z^n} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{nz^n}{1-z^n}.$$

It seems probable that he made a complete study of the more general series now known as the Lambert series:

$$L(z) = \sum_{n=1}^{\infty} a_n \frac{z^n}{1-z^n}.$$

About a century later Weierstrass[3] treated the series

$$\sum_{n=1}^{\infty} \frac{z^n}{1+z^{2n}}.$$

He showed that its region of convergence consists of two distinct parts, the interior and the exterior of the unit circle, in each of which the series represents a function which cannot be continued over the boundary $|z| = 1$.

In 1907 Hansen[4] set up the series

$$\sum_{n=1}^{\infty} \frac{z^{rn+t}}{1-z^{rn+t}}.$$

It reduces to the Lambert series with coefficient unity when $r = 1$ and $t = 0$, while the difference of two series of this type gives that of Weierstrass.

Knopp, in 1913, published an extensive treatment of the Lambert series, in which he discusses its region of convergence and shows that under certain restrictions placed on the coefficients a_n, the function represented by the series cannot be continued beyond the unit circle.[5]

---

[1] This paper was presented to the American Mathematical Society on June 20, 1934 under the title, "On the convergence of a generalized series and the relation of its coefficients to those of the corresponding power series."

[2] J. H. Lambert, *Anlage zur Architektonik*, vol. 2, Riga, 1771.

[3] L. Weierstrass, "Zur Functionentheorie," *Monatsberichte der Berliner Akademie* (1880), pp. 719-743.

[4] C. Hansen, "Démonstration de l'impossibilité du prolongement analytique de la série de Lambert," *Oversigt over det kgl. danske videnskabernes selskabs fordhandlinger* (1907), pp. 1-19.

[5] K. Knopp, "Ueber Lambertsche Reihen," *Journal für Mathematik*, vol. 142 (1913), pp. 283-315.

The purpose of this paper is to present a series which includes the three above mentioned as particular cases and to show how some remarkable properties of these series can be extended to a much larger class.

II. *Definition of the series; its convergence.* The series which we shall study is:

$$\sum_{n=1}^{\infty} a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} \tag{2.1}$$

where λ and μ are any positive integers. We shall refer to it briefly as the F-series. It reduces to the Lambert series $L(z)$ when $\lambda = \mu = 1$, or to $L(z^k)$ when $\lambda = \mu = k$. The series of Weierstrass is the difference of two series of type (2.1):

$$\sum_{n=1}^{\infty} \frac{z^n}{1 + z^{2n}} = \sum_{n=1}^{\infty} \frac{z^n}{1 - z^{4n}} - \sum_{n=1}^{\infty} \frac{z^{3n}}{1 - z^{4n}}. \tag{2.2}$$

Finally, if a_n is any constant c, we obtain

$$\sum_{n=1}^{\infty} \frac{cz^{\lambda n}}{1 - z^{\mu n}} = \sum_{n=1}^{\infty} \frac{cz^{\mu n + (\lambda - \mu)}}{1 - z^{\mu n + (\lambda - \mu)}}. \tag{2.3}$$

This is Hansen's series if $c = 1$.

It is important at the outset to determine the exact region of convergence of the F-series, and for this we have the following

THEOREM 1.[6] (A) *If Σa_n diverges, and* (i) $\mu > \lambda$, *the F-series and the power series $\Sigma a_n z^{\lambda n}$ converge and diverge at the same points within the unit circle, while for values of z beyond the unit circle the F-series and the power series $\Sigma a_n z^{(\lambda-\mu)n}$ converge and diverge together.*

If Σa_n diverges and (ii) $\mu \leq \lambda$, *the F-series and $\Sigma a_n z^{\lambda n}$ converge and diverge at the same points when $|z| < 1$, but the F-series diverges for every $|z| > 1$.*

(B) *If Σa_n converges and* (i) $\mu \geq \lambda$, *the F-series converges for every z whose modulus is not equal to unity.*

If Σa_n converges and (ii) $\mu < \lambda$, *the F-series converges for all $|z| < 1$, but converges and diverges at the same points as $\Sigma a_n z^{(\lambda-\mu)n}$ when $|z| > 1$.*[7]

The proof of this theorem is left to the reader. It is easily shown that its results can be extended to series in which $\lambda \leq 0$ or $\mu < 0$ or both.[8]

---

[6] Throughout the remainder of this paper all summations, unless otherwise specified, are to be taken from $n = 1$ to $n = \infty$.

[7] This theorem does not exclude the possibility of convergence at some points on the unit circle, but such cases will not be considered.

[8] In establishing further properties of the series it will not be necessary to take into

By means of Weierstrass's M-test the uniform convergence of the series is established, so that we have

THEOREM 2. *The F-series converges uniformly in every closed sub-region lying completely within one of its regions of convergence and including no point of modulus unity.*

Whenever the F-series converges for values of z such that $|z| > 1$, a relation can be set up between the sum of the series for a point z outside the unit circle and the sum of some typical F-series at a point $1/z$ inside it. If $\lambda < \mu$ we have:

$$\Sigma a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} = -\Sigma a_n \frac{(1/z)^{(\mu-\lambda)n}}{1-(1/z)^{\mu n}}, \tag{2.4}$$

while for $\mu = \lambda$ this becomes:

$$\Sigma a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} = -\Sigma a_n - \Sigma a_n \frac{(1/z)^{\mu n}}{1-(1/z)^{\mu n}}. \tag{2.5}$$

When $\lambda > \mu$ we can use a lemma similar to that developed by Ananda-Rau [9] for the Lambert series, and obtain the relation:

$$\Sigma a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} = -\Sigma a_n \frac{(1/z)^{(q\mu-k)n}}{1-(1/z)^{\mu n}} - \sum_{p=1}^{q} f[(1/z)^{(p-1)\mu-k}], \tag{2.6}$$

where $k = \lambda - \mu$, $f(z^s) = \Sigma a_n z^{sn}$, and q is the smallest positive integer for which $q\mu > k$.

Hence we may limit our considerations to the region of convergence of the F-series which lies within the unit circle. The radius of this region is $r \leq 1$.

III. *Expansion as a power series; inversion of a power series.* Every term of the F-series is analytic in $|z| < r$ and the series is uniformly convergent in $|z| \leq \rho < r$. Therefore the theorem of Weierstrass [10] on double series may be applied to obtain the expansion in power series of the function represented by the F-series in $|z| < r$. The result is:

$$\Sigma a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} = \Sigma A_n z^n, \tag{3.1}$$

---

consideration any but positive values of λ and μ since a series in which $\lambda \leq 0$, or $\mu < 0$ can be transformed in such a way that the theorems developed for positive values will be applicable.

[9] K. Ananda-Rau, "On Lambert's series," *Proceedings of the London Mathematical Society*, ser. 2, vol. 18 (1920), p. 3.

[10] K. Knopp, *Theory and Application of Infinite Series*, Blackie and Son, Ltd., London and Glasgow, 1928, p. 430.

where A_n is the sum of all the coefficients a_k whose subscript k is such a divisor of n that $n \equiv \lambda k \pmod{\mu k}$, or, to borrow a notation from Knopp, $A_n = \sum_{k|n}' a_k$.

In order that a power series may be expressed as an F-series it is necessary and sufficient that the coefficients a_n be known in terms of the A's appearing in the power series. If we consider the case in which λ and μ are relatively prime or are multiples of relatively prime integers, we shall find by expanding the series that the a_n's of the F-series are not in general uniquely expressible in terms of A's. Similar results are obtained when λ is a multiple of μ. But if $\mu = m\lambda$ $(m = 1, 2, 3, \cdots)$ every a_n is a unique sum of certain A's in the power series $\Sigma A_{\lambda n} z^{\lambda n}$. Thus we have

THEOREM 3. *The necessary and sufficient condition that a power series* $\Sigma A_{\lambda n} z^{\lambda n}$ *be expressible as an F-series is that* μ *be a multiple of* λ.

To determine the form of the relation which expresses any given a_n in terms of A's, we shall introduce an inversion function denoted by $R(n)$. Let $\{n\}$ be a set of positive integers defined by $n \equiv 1 \pmod{m}$, where $m = 1, 2, 3, \cdots$, each value of m determining a particular set $\{n\}$. We shall call the k-divisors of any integer n those divisors which satisfy the congruence $n/k \equiv 1 \pmod{m}$. For the elements of the set $\{n\}$, $R(n)$ is defined to be an integer, positive or negative, such that

$$(3.2) \qquad \sum_{k|n} R(k) = R(1) + R(k_1) + R(k_2) + \cdots + R(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } n > 1, \end{cases}$$

where $1, k_1, k_2, \cdots, n$ are the k-divisors of n.

When $m = 1$, $R(n)$ reduces to the Möbius function $\mu(n)$,[11] for then $\{n\}$ is the complete set of positive integers and the k-divisors are all the factors of n. When $m = 2$, $R(n)$ may be defined similarly to the Möbius function, $\{n\}$ being the set of all the odd positive integers whose k-divisors are the odd factors of n.

We can now show that if $A_{\lambda n} = \sum_{k|n} a_k$, the inverse relation is given by

$$a_n = \sum_{k|n} R(n/k) A_{\lambda k}.$$

Since $A_{\lambda k} = \sum_{\kappa|k} a_\kappa$, it follows that

$$\sum_{k|n} R(n/k) A_{\lambda k} = \sum_{k|n} R(n/k) \sum_{\kappa|k} a_\kappa.$$

[11] A. F. Möbius, "Ueber eine besondere Art von Umkehrung der Reihen," *Journal für Mathematik*, vol. 9 (1832), p. 105.

For a fixed κ, the coefficient of a_κ is

$$\sum_{K|n/\kappa} R\left(\frac{n/\kappa}{K}\right) = \sum_{K|n/\kappa} R(K),$$

since a given a_κ will have as coefficient only those terms of the sum $\sum_{k|n} R(n/k)$ for which k has the factor κ, that is, $k = \kappa K$. But by the definition of $R(n)$, the sum $\sum_{K|n/\kappa} R(K)$ vanishes unless $n/\kappa = 1$, that is, $n = \kappa$, in which case the coefficient of a_n is unity, since $\sum_{K|1} R(K) = R(1) = 1$. Hence

$$\sum_{k|n} R(n/k)\ A_{\lambda k} = a_n. \tag{3.3}$$

Thus any power series $\Sigma A_{\lambda n} z^{\lambda n}$ can be expressed as an F-series in which μ is a multiple of λ and a_n is given by (3.3).

IV. *Existence of natural boundaries.* Knopp has shown that when $a_n = 1$ in the Lambert series the function represented by the series cannot be continued beyond the unit circle. This property can be extended to series having $\mu = \lambda = k$ and also to those in which $\mu \gtrless \lambda$, as given in the following theorem, the method of proof being that of Landau as presented by Knopp.[12]

THEOREM 4. *If in the F-series $a_n = 1$ and* (i) $\mu \leq \lambda$, *the series represents a function $F(z)$ which cannot be continued beyond the unit circle;* (ii) $\mu > \lambda$, *it represents two functions in $|z| < 1$ and $|z| > 1$ respectively, both of which have the unit circle as a natural boundary.*

The proof of this theorem consists in showing that $F(z)$ becomes infinite as $z \to z_0$, z_0 being one of a set of points everywhere dense on the unit circle. If we take z_0 to be a rational point, namely, $z_0 = e^{2\pi i l'/l}$, where l is any positive integer except one of the finite number of factors of μ, and l' is relatively prime to l, Landau's method can be applied.

For values of a_n other than unity, Knopp[13] has demonstrated that under certain restrictions placed upon a_n, Lambert's series represents a function for which $|z| = 1$ is a natural boundary, and a parallel to his theorem can be developed for the F-series.

THEOREM 5. *If the coefficients a_n of the F-series are such that the radius of convergence of $\Sigma a_n z^{\lambda n}$ is unity, and also such that for an integer l (except the finite number of factors of μ) the series*

[12] K. Knopp, *Ueber Lambertsche Reihen,* p. 291.
[13] *Ibid.*, p. 292.

$$\sum_{\nu=1}^{\infty} \frac{a_{l\nu+q}}{\lambda\nu+q} \qquad [q=0,1,2,\cdots,(l-1)]$$

converges, and if for such an l and an l' relatively prime to l, $z_0 = e^{2\pi i l'/l}$, then for all positive values of λ and μ

$$\lim_{z\to z_0}\left\{\left(1-\frac{z}{z_0}\right)\Sigma a_n \frac{z^{\lambda n}}{1-z^{\mu n}}\right\} = \frac{1}{\mu}\sum_{\nu=1}^{\infty}\frac{a_{l\nu}}{l\nu}.$$

If this be true for an infinite number of l's and if $\sum_{\nu=1}^{\nu=\infty} a_{l\nu}/l\nu \neq 0$, then the function represented by the F-series cannot be continued beyond the unit circle.

Proof. We wish to determine

$$\lim_{z\to z_0}\left(1-\frac{z}{z_0}\right)\Sigma a_n\frac{z^{\lambda n}}{1-z^{\mu n}} \quad \text{or} \quad \lim_{\rho\to 1}(1-\rho)\Sigma a_n\frac{z^{\lambda n}}{1-z^{\mu n}}.$$

Following Knopp's procedure at this point, let us break up the series into Σ_1 and Σ_2 according as $n \equiv 0$ or $\not\equiv 0 \pmod{l}$. Then if $n = l\nu$,

$$\begin{aligned}
\lim_{\rho\to 1}(1-\rho)\Sigma_1 &= \lim_{\rho\to 1}(1-\rho)\sum_{\nu=1}^{\infty} a_{l\nu}\frac{\rho^{\lambda l\nu}}{1-\rho^{\mu l\nu}}\\
&= \lim_{\rho\to 1}\frac{1-\rho}{1-\rho^{\mu l}}(1-\rho^{\mu l})\sum_{\nu=1}^{\infty} a_{l\nu}\frac{\rho^{\lambda l\nu}}{1-\rho^{\mu l\nu}}\\
&= \frac{1}{\mu l}\lim_{y\to 1}(1-y)\sum_{\nu=1}^{\infty} a_{l\nu}\frac{y^{\lambda\nu/\mu}}{1-y^{\nu}} \qquad (y=\rho^{\mu l})\\
&= \frac{1}{\mu}\lim_{y\to 1}\sum_{\nu=1}^{\infty}\frac{a_{l\nu}}{l\nu}\,\frac{\nu y^{l\nu/\mu}}{1+y+y^2+\cdots+y^{\nu-1}}.
\end{aligned}$$

Since $\sum_{\nu=1}^{\nu=\infty} a_{l\nu}/l\nu$ is convergent, it follows that the entire series on the right is uniformly convergent in $0 \leq y \leq 1$ if $\nu y^{\lambda\nu/\mu}/(1+y+y^2+\cdots+y^{\nu-1})$ is uniformly bounded for every $0 \leq y \leq 1$ and for every $\nu \geq 1$.[14] Now for $0 < y < 1$ and for $\nu \geq 1$

$$\begin{aligned}
\frac{\nu y^{\lambda\nu/\mu}}{1+y+y^2+\cdots+y^{\nu-1}} = \frac{\nu y^{\lambda\nu/\mu}(1-y)}{1-y^{\nu}} &\leq \frac{\nu y^{\lambda\nu/\mu}(1-y)}{1-y}\\
&= \nu y^{\lambda\nu/\mu} < k,
\end{aligned}$$

since $\nu(y^{\lambda/\mu})^{\nu}$ is a null sequence and therefore bounded. For $y = 0$,

$$\frac{\nu y^{\lambda\nu/\mu}}{1+y+y^2+\cdots+y^{\nu-1}} = 0$$

[14] K. Knopp, *Theory and Application of Infinite Series*, p. 346.

and for $y = 1$ it is equal to unity. Thus for every $0 \leqq y \leqq 1$ and for every $\nu \geqq 1$

$$\frac{\nu y^{\lambda\nu/\mu}}{1 + y + y^2 + \cdots + y^{\nu-1}} < K,$$

where K is the larger of the two quantities 1 and k. Hence the series is uniformly convergent in $0 \leqq y \leqq 1$ and

$$\lim_{\rho\to 1} (1-\rho)\Sigma_1 = \frac{1}{\mu} \lim_{y\to 1} \sum_{\nu=1}^{\infty} \frac{a_{l\nu}}{l\nu} \frac{\nu y^{\lambda\nu/\mu}}{1 + y + y^2 + \cdots + y^{\nu-1}}$$
$$= \frac{1}{\mu} \sum_{\nu=1}^{\infty} \lim_{y\to 1} \frac{a_{l\nu}}{l\nu} \frac{\nu y^{\lambda\nu/\mu}}{1 + y + y^2 + \cdots + y^{\nu-1}} = \frac{1}{\mu} \sum_{\nu=1}^{\infty} \frac{a_{l\nu}}{l\nu}.$$

Knopp's method may be applied to show that

$$\lim_{z\to z_0} \left\{ \left(1 - \frac{z}{z_0}\right) \Sigma_2 \right\} = 0,$$

or that for every $q = 1, 2, \cdots, (l-1)$

$$\lim_{z\to z_0} \left\{ \left(1 - \frac{z}{z_0}\right) \sum_{\nu=0}^{\infty} a_{l\nu+q} \frac{z^{\lambda(l\nu+q)}}{1 - z^{\mu(l\nu+q)}} \right\} = 0.$$

Thus we have

$$\lim_{z\to z_0} \left\{ \left(1 - \frac{z}{z_0}\right) \Sigma a_n \frac{z^{\lambda n}}{1 - z^{\mu n}} \right\} = \frac{1}{\mu} \sum_{\nu=1}^{\infty} \frac{a_{l\nu}}{l\nu}.$$

If this be true for an infinite number of l's and $\sum_{\nu=1}^{\nu=\infty} a_{l\nu}/l\nu \neq 0$, then all the rational points defined by z_0 are singular points and the function represented by the F-series cannot be continued beyond the unit circle.

We have taken into consideration the radial approach of z to z_0 from within the unit circle, and have therefore shown that when the F-series converges for $|z| < 1$ the function which it represents has the unit circle as a natural boundary. However, whenever the F-series converges for $|z| > 1$ it can be expressed in terms of some other F-series convergent for $|z| < 1$, to which the theorem applies. Thus, if the series converges for $|z| < 1$ and also for values of z such that $|z| > 1$, it will in general represent two functions, neither of which can be continued across the unit circle.

Notre Dame College,
South Euclid, Ohio.

GEODESIC CONTINUA IN ABSTRACT METRIC SPACE.

By Orrin Frink, Jr.

Menger has treated the subject of geodesics in abstract metric space. (*Mathematische Annalen*, vol. 103 (1930), pp. 466-501.) In Menger's work, a geodesic is an arc whose length, properly defined, is less than or equal to the lengths of comparison arcs. Existence theorems for geodesics, with more general end conditions than those of Menger, can be proved by showing that the length or linear measure function is lower semicontinuous.

This method of approach to the subject of geodesic arcs by means of lower semicontinuity is subject to the difficulty that the limit set of a convergent sequence of arcs is not necessarily an arc. Such a limit set of arcs is, however, always a continuum, which suggests generalizing the notion of a geodesic arc to that of geodesic continuum. A geodesic continuum may be defined as one whose length, or linear measure, is less than or equal to the linear measures of comparison continua.

Menger has shown (*loc. cit.*, p. 477) that one of his definitions of length, called "Längeninhalt," which is defined for continua as well as for arcs, is lower semicontinuous. However, this particular definition of length, while satisfactory for arcs, has disadvantages when applied to continua in general. For, as can be seen from an example given by Menger (*loc. cit.*, p. 476), the Längeninhalt of a set consisting of the two diagonals of a unit square is 3, while each diagonal by itself has a Längeninhalt equal to $\sqrt{2}$. Menger does not use his theorem on lower semicontinuity in proving his existence theorems.

In the present paper a form of the Caratheodory definition of linear measure, applicable to metric sets in general, is used, rather than any special definition of length suitable only for arcs. It is shown that for continua, the Caratheodory linear measure is lower semicontinuous. This, taken together with some results on the compactness of collections of continua, leads to existence theorems for geodesic continua with rather general end conditions. (See Theorems 4, 5, and 7 below.) Aside from the gain in generality due to the end conditions, it is interesting to note that many of the functionals of the calculus of variations, defined in terms of an integral J, become cases of the Caratheodory linear measure function when the metric of the space is redefined in terms of J. (Morse, *Calculus of Variations in the Large*, p. 208; Menger, *Fundamenta Mathematica*, vol. 25, p. 441.)

Definitions.

Two sets A and B are said to be *adjacent* if they have a point in common, or if either contains a limit point of the other. Note that a set is connected if it is not the sum of two non-empty sets which are not adjacent.

The *distance* $\overline{AB}$ between two closed sets A and B is the greatest lower bound of numbers ϵ such that $A \subset U(B, \epsilon)$ and $B \subset U(A, \epsilon)$, where $U(A, \epsilon)$ means the ϵ-neighborhood of A. (Hausdorff, *Mengenlehre*, 2nd ed., p. 146.)

An *ϵ-partition* of a set A is a collection of a finite number of subsets $A_1, A_2, \cdots, A_n$ of A, whose logical sum is A, the diameter $d(A_r)$ of each subset being less than ϵ. The subsets $A_1, \cdots, A_n$ are called the *subdivisions* of the partition, and their number, n, is called the *degree* of the partition. The sum of the diameters of all the subdivisions is called the *diameter sum* of the partition, and is less than $n\epsilon$.

The (Caratheodory) *linear measure* $m(A)$ of the set A is the greatest lower bound of numbers l such that for every $\epsilon > 0$ there exists an ϵ-partition of A whose diameter sum is $< l$. If there is no such number l, then $m(A)$ is said to be infinite. (Caratheodory, *Göttinger Nachrichten* 1914, pp. 404-426.) A more general form of the Caratheodory definition, where the number of subdivisions is allowed to be countably infinite, is not needed for continua, with which this paper is chiefly concerned.

Lemmas.

The first five lemmas are stated without proof.

Lemma 1. *Every subdivision of a partition of a connected set is adjacent to at least one other subdivision of the partition.*

Lemma 2. *If A and B are adjacent, then $d(A + B) \leqq d(A) + d(B)$.*

Lemma 3. *The diameter sum of a partition is not increased if the partition is altered by replacing two adjacent subdivisions by their logical sum.*

Lemma 4. $d[U(A, \epsilon)] \leqq d(A) + 2\epsilon$.

Lemma 5. *If the sets $B_1, B_2, \cdots, B_n$ are each adjacent to A, and are each of diameter less than ϵ, then*

$$d(A + B_1 + B_2 + \cdots + B_n) < d(A) + 2\epsilon.$$

The next lemma states a property of partitions of connected sets that does not hold for disconnected sets. This accounts for the fact that the linear measure function is not a lower semicontinuous function of sets in general, but only of connected sets.

LEMMA 6. *If A is connected, and $m(A) < l$, then for every $\epsilon > 0$ there exists an ϵ-partition of A of diameter sum $< l$, every subdivision of which is of diameter $\geqq \epsilon/4$.*

Proof. Since $m(A) < l$, by definition of linear measure there exists an $\epsilon/2$-partition of A of diameter sum $< l$. The idea of the proof is to combine adjacent subdivisions of this partition, using Lemma 3, until all subdivisions are of diameter $\geqq \epsilon/4$. Accordingly, let two adjacent subdivisions of diameter $< \epsilon/4$ be replaced by their logical sum, and let this step be repeated until no two adjacent subdivisions are each of diameter $< \epsilon/4$. It follows from Lemmas 2 and 3 that the new partition is still an $\epsilon/2$-partition of diameter sum $< l$.

The new partition may still have subdivisions of diameter $< \epsilon/4$, but by Lemma 1 each such subdivision must be adjacent to one of diameter $\geqq \epsilon/4$. The next step is to replace such a subdivision of diameter $\geqq \epsilon/4$, together with all subdivisions of diameter $< \epsilon/4$ adjacent to it, by the logical sum of all these subdivisions. This step is repeated until no subdivisions of diameter $< \epsilon/4$ remain, which is one of the things to be proved. The resulting partition is an ϵ-partition by Lemma 5, although it may no longer be an $\epsilon/2$-partition. By Lemma 3, its diameter sum is still $< l$. This completes the proof of Lemma 6.

LEMMA 7. *If A is a connected set, and $m(A) < l$, then for every $\epsilon > 0$ there exists an ϵ-partition of A whose diameter sum is $< l$, and whose degree is $< 4l/\epsilon$.*

This follows from Lemma 6. For if the partition provided by Lemma 6 were of degree $\geqq 4l/\epsilon$, then since the diameter of each subdivision is $\geqq \epsilon/4$, the diameter sum of the partition would be $\geqq 4l/\epsilon \cdot \epsilon/4 = l$, contrary to Lemma 6.

THEOREMS ON LOWER SEMICONTINUITY.

THEOREM 1. *In a metric space, if A is any set of linear measure $\geqq l$, then for every $\theta > 0$ there exists a $\delta > 0$ such that if C is any connected set for which $A \subset U(C, \delta)$, then $m(C) > l - \theta$.*

This theorem states that connected sets sufficiently close to a fixed set have a linear measure not much less than that of the fixed set.

Proof. Since $m(A) \geqq l$, it follows from the definition of linear measure that for $\theta > 0$ there exists an $\epsilon > 0$ such that no ϵ-partition of A has a diameter sum $< l - \theta/3$. For otherwise $m(A)$ would be $\leqq l - \theta/3$.

Now choose the δ required by the theorem such that $\delta < \epsilon/4$ and $\delta < (\epsilon \cdot \theta)/48l$. For this choice of δ it must be shown that if C is any connected set such that $A \subset U(C, \delta)$, then $m(C) > l - \theta$. Suppose on the contrary that $m(C) \leqq l - \theta$. Then by Lemma 7 there exists an $(\epsilon/2)$-partition of C whose diameter sum is $< l - 2\theta/3$, and whose degree is $< 8l/\epsilon$. Let $\{C_k\}$ be the subdivisions of this partition, and n its degree.

By hypothesis $A \subset U(C, \delta)$, hence the sets $\{A \cdot U(C_k, \delta)\}$ form the subdivisions of a partition of A, also of degree n. To estimate the diameter sum of this ϵ-partition of A, note that from Lemma 4 it follows for any particular subdivision $A \cdot U(C_k, \delta)$ of this partition that $d[A \cdot U(C_k, \delta)] \leqq d(C_k) + 2\delta$. Hence the diameter sum of this partition of A exceeds the diameter sum of the partition of C by at most $2n\delta$, there being n subdivisions. Since $n < 8l/\epsilon$, and $\delta < (\epsilon \cdot \theta)/48l$, it follows that $2n\delta < \theta/3$.

That is, the diameter sum of the partition of A exceeds the diameter sum of the partition of C by less than $\theta/3$. But the diameter sum of the partition of C was $< l - 2\theta/3$. Hence the diameter sum of the partition of A is $< l - \theta/3$, contrary to the assumption above that no such partition exists. This contradiction proves Theorem 1.

THEOREM 2. *In a metric space whose elements are continua of some other metric space, the Caratheodory linear measure function is lower semi-continuous at every element of the space.*

Proof. To prove lower semicontinuity at an element A it must be shown that if $m(A) \geqq l$, then for every $\theta > 0$ there exists a $\delta > 0$ such that if C is another element (continuum) of the space, for which the distance $\overline{AC} < \delta$, then $m(C) > l - \theta$. Theorem 1 supplies just such a value of δ, since if $\overline{AC} < \delta$, by definition of the distance $\overline{AC}$ it follows that $A \subset U(C, \delta)$, which is the condition of Theorem 1. Since the sets A and C are continua, C is connected, and both A and C are closed, so that the definition of distance between closed sets applies.

APPLICATIONS OF LOWER SEMICONTINUITY.

To derive an existence theorem for geodesic continua from Theorem 2, it is sufficient to show that the collection of admissible continua, among which one of minimum linear measure is to be found, is closed and compact, in which case the theorem applies that a function lower semicontinuous on a closed and compact set attains its minimum value for some element of the set.

Hausdorff has shown (*Grundzüge der Mengenlehre*, 1st ed., p. 302; *Mengenlehre*, 2nd ed., p. 150) that the space of all continua of a compact

metric space is itself a compact metric space. Hence the collection of all continua of a compact metric space which satisfy given end conditions is closed and compact provided the limit element of any convergent sequence of the collection also satisfies the end conditions.

THEOREM 3. *In a metric space, if the closed set A is the limit of the sequence of closed sets $\{C_r\}$, and every set C_r of the sequence has at least one point in common with the closed and compact set B, then A has at least one point in common with B.*

Proof. Suppose A and B have no common point. Then, since A and B are closed, and B is compact, there exists a δ-neighborhood $U(A, \delta)$ of A which contains no point of B. Since A is the limit of the sequence $\{C_r\}$, there exists a C_r such that $C_r \subset U(A, \delta)$. But this is impossible, since C_r contains a point of B, and $U(A, \delta)$ does not. This proves Theorem 3.

COROLLARY. *All continua of a compact metric space which have at least one point in common with each set of a collection of closed sets, constitute a compact metric space.*

This follows from Theorem 3. For, the collection of all continua having at least one point in common with one fixed closed set C_r of the collection of closed sets $\{C_r\}$, is closed by Theorem 3. Hence the logical product of all such collections is closed. But this logical product is just the collection of continua mentioned in the corollary. Being a closed subset of the compact metric space of all continua, its elements constitute a compact metric space. Similar results may be found in R. L. Moore's *Foundations of Point Set Theory*, Chapter 5.

As a consequence of the corollary, together with Theorem 2 and the property that a function which is lower semicontinuous over a compact space assumes its minimum, we have

THEOREM 4. *In a compact metric space, if there exists a continuum of finite linear measure which has at least one point in common with each set of a collection of closed sets, then there exists a continuum of minimum linear measure having this property.*

Each closed set of the collection mentioned in Theorem 4 corresponds to an end condition on the geodesic continuum. If the closed set consists of a single point, it is a fixed end point condition; otherwise a variable end point condition. In particular, if each closed set of the collection consists of a single point, the end conditions amount to requiring that the continua being con-

sidered all contain the set A (consisting of all these single points). The set A is arbitrary. This proves

THEOREM 5. *In a compact metric space, if there exists a continuum of finite linear measure containing the set A, then there exists a continuum of minimum linear measure containing A.*

A more general type of variable end condition, corresponding to the end conditions of Morse, *The Calculus of Variations in the Large*, pp. 19 and 65, is obtained if in addition to the condition that the continuum have a point, say c_r, in common with each set C_r of a collection $\{C_r\}$ of closed sets, it is required that certain relations hold between the endpoints $\{c_r\}$. Such relations between the endpoints are most conveniently described in terms of the notion of the product space P of the closed end sets $\{C_r\}$.

Suppose in the compact metric space M there are given n such end sets $C_1, \cdots, C_r, \cdots, C_n$, no two having a point in common. To every selection $(c_1, \cdots, c_r, \cdots, c_n)$ of points, one from each of the end sets C_r, corresponds a point π of the product space P. The points $(c_1, \cdots, c_n)$ corresponding to π will be called the *coördinates* of π. The notions of neighborhood and distance in the product space are defined in the usual way.

Now let Q be any closed set of the product space P. A continuum K of the original compact metric space M is said to satisfy the *end condition* Q if K contains all the coördinates $(c_1, \cdots, c_n)$ of at least one point π of Q. This type of end condition is stronger than that of Theorem 4, since K is still required to have at least one point in common with each end set, and in addition the end points must satisfy the condition of being the coördinates of some point π of Q. What is now needed is the result that the collection of all continua satisfying an end condition Q is closed and compact.

THEOREM 6. *If Q is a closed set of the product space P of n mutually exclusive closed sets $C_1, \cdots, C_n$ of the compact metric space M, then the space of all continua of M which contain all the coördinates $(c_1, \cdots, c_n)$ of at least one point π of Q, is a compact metric space.*

Proof. Suppose the continuum K is the limit of the convergent sequence of continua $\{K_m\}$, and suppose each continuum K_m of the sequence satisfies the end condition Q. It must be shown that K also satisfies the end condition Q. For each m, let π_m be the point of Q corresponding to K_m. Then K_m contains all the coördinates of π_m. Since Q is closed and compact, the sequence $\{\pi_m\}$ contains a subsequence $\{\pi_{m_j}\}$ converging to a point π of Q. It will now be shown that K contains all the coördinates of π. Let c_p be the

p-th coördinate of π, and c_{pj} be the p-th coördinate of π_{mj}. Since the sequence $\{\pi_{mj}\}$ converges to π, it follows that for every p the sequence $\{c_{pj}\}$ converges to c_p. Now since K is the limit of the convergent sequence of continua $\{K_{mj}\}$, every δ-neighborhood of K contains all except a finite number of the continua $\{K_{mj}\}$, and hence every δ-neighborhood of K contains all except a finite number of the points $\{c_{pj}\}$. Since K is closed, it follows that K also contains the limit c_p of the sequence $\{c_{pj}\}$. This is true for every p, hence K contains all the coördinates $\{c_p\}$ of π, and therefore satisfies the end condition Q. This proves that the collection of all continua of M which satisfy the end condition Q, is closed, since any limit continuum of the collection also satisfies the end condition Q, and is therefore a member of the collection. Since the collection is a closed subset of the compact metric space of all continua of M, it is itself a compact metric space.

THEOREM 7. *If Q is a closed set of the product space P of n mutually exclusive closed sets $(C_1, \cdots, C_n)$ of the compact metric space M, and if there exists a continuum of M of finite linear measure which contains all the coördinates of at least one point of Q, then there exists a continuum of minimum linear measure having this property.*

This follows from Theorems 2 and 6.

In the existence Theorems 4 and 7, if there are more than two end sets C_r, it cannot be expected in general that the geodesic continuum whose existence is asserted will be an arc. However, it is interesting to note that in the important case where there are just two end sets C_1 and C_2, the geodesic continuum must be an arc. To show this, it is sufficient to show that the geodesic continuum is locally connected and irreducible between its endpoints (Hausdorff, *Mengenlehre*, 2nd ed., p. 222). It follows from a theorem of R. L. Moore (*Foundations of Point Set Theory*, p. 95, Theorem 8) that a continuum of finite linear measure is locally connected. For Moore's theorem asserts that if the continuum K is not locally connected, then K contains infinitely many mutually exclusive subcontinua, each of diameter greater than a fixed constant. From this it follows that the linear measure of K is infinite. That a geodesic continuum with two end conditions is irreducible between its endpoints follows from the fact that it is geodesic. For, any proper subcontinuum of K is necessarily of smaller linear measure than K. This proves

THEOREM 8. *A geodesic continuum with two end conditions is an arc.*

THE PENNSYLVANIA STATE COLLEGE,
STATE COLLEGE, PA.

PROOF OF THE IDEAL WARING THEOREM FOR EXPONENTS 7-180.

By L. E. DICKSON.

1. If q denotes the greatest integer $< (3/2)^n$, then $2^n q - 1$ is a sum of $I = 2^n + q - 2$, but not fewer, n-th powers (Lemma 1). The ideal Waring theorem states that every positive integer is a sum of I n-th powers; for example, 4 squares, 9 cubes, 19 fourth powers. Proofs for squares and cubes are classic. For $n \leqq 20$, proof has been found, but not yet published, by use of a "constant" far exceeding our new constant N in Theorem 1, and the use of more or less extensive tables.

Using the new N, we here prove without any tables the ideal Waring theorem for $9 \leqq n \leqq 180$, and by use of tables also for $n = 7$ and 8. For $n = 6$ we cannot attain the ideal 73, but reach 115 (the best earlier result being 160).

PART I. ASYMPTOTIC THEORY.[1]

2. If $(x]$ denotes the distance from x to the nearest integer, let

$$\Omega' \leqq \sum_{y=1}^{Y} \sum_{y_1=1}^{Y} \frac{1}{2(\alpha t]}, \qquad t = y^n - y_1^n, \quad t \neq 0.$$

This sum is double of that obtained when $t > 0$. When t and y_1 are fixed there is a single real positive y. Thus $y^n - y_1^n$ has at most Y sets of integral solutions y, y_1, each chosen from $1, \cdots, Y$. Hence

$$\Omega' \leqq Y \Sigma 1/(\alpha t],$$

summed for certain integers $t \geqq 1$, $t \leqq Y^n$. Let l be the absolutely least residue of at modulo q. By A, § 13, $1/(\alpha t] < 2q/i$, where $i = | l |$ and the values of i are in $(1, 1)$ correspondence with those of $t > 0$. In ascending order of magnitude, the values of i are $\geqq 1, \geqq 2, \cdots$, terminating at or before Y^n. Hence

$$\Omega' \leqq 2Yq \sum_{j=1}^{Y^n} (1/j) \leqq 2Yq(1 + \log Y^n),$$

as is shown by comparing the sum (for $j \geqq 2$) of the rectangles under the curve $y = 1/x$ with the area $=$ integral from 1 to Y^n.

[1] We avoid the divisor function in § 13 of our exposition and generalization of Vinogradow's theory, *Annals of Mathematics*, vol. 36 (1935), pp. 395-405, cited as A.

Since we are dealing with intervals of the second class, $q < \tau = 2mnP^{n-\frac{1}{2}}$. Hence (92) of A becomes

$$|\sum_{Y} V_y S_y|^2 < X_1 R Y(a 2^n R^n + 4mnP^{n-\frac{1}{2}}(1 + \log Y^n)),$$

in which we have dropped the subscript $k + k_1$ of a. By (9), A,

$$(1) \qquad Y^n \leqq CP^{(n-1)f}, \qquad C = m/(a2^{n-1}),$$

$$(2) \qquad 1 + \log Y^n \leqq 1 + \log C + (n-1)f \log P$$

will be later verified to be $\leqq P^z$. By (7), A, $R \leqq P^{1-f}$. Hence

$$\begin{aligned} |\Sigma V_y S_y|^2 &< X_1 R Y\{a 2^n P^{n(1-f)} + 4mnP^{z+n-\frac{1}{2}}\} \\ &< X_1 R Y \cdot 8mnP^{z+n-\frac{1}{2}}, \end{aligned}$$

since the latter exponent of P exceeds $n(1-f)$ by more than z in view of $f \geqq \frac{1}{2}\nu$, $\nu = 1/n$. Hence by A, (103), and the line below it, we see that (88) gives

$$(3) \qquad |H_2| < DP^{s-2}PX(X_1 RYP^{z+n-\frac{1}{2}})^{\frac{1}{2}}, \qquad D = 3(8mn)^{\frac{1}{2}},$$

By (A), (7), (9), (19), (24),

$$(4) \qquad R > \sqrt{1/2}\, P^{1-f}, \quad Y > \sqrt{1/2}\, C^\nu P^{(1-\nu)f}, \quad X \geqq \tfrac{1}{2}\kappa n^\mu P^{n-1-\sigma},$$

while X_1 exceeds the last product with P replaced by R, and κ by κ_1, since we assume that $k_1 = k$. Note that

$$(5) \qquad \mu = \sigma - n + 1, \qquad \sigma = n(1-\nu)^k.$$

Multiply and divide (3) by $X(X_1 RY)^{\frac{1}{2}}$ and apply (4). Thus

$$(6) \qquad |H_2| < C_1 P^J X^2 X_1 R Y P^{s-n}, \qquad C_1 = \frac{4D 2^{(n-1-\sigma)/4}}{\kappa n^\mu (\kappa_1 n^\mu C^\nu)^{1/2}},$$

$$(7) \qquad 2J = \sigma(3-f) - g, \;\; g = \tfrac{1}{2} - z - nf + (1-\nu)f.$$

The Waring theorem is true for every integer $\geqq N_0$, if the integral $I(N_0) > 0$. By the above and A, (89), the condition is

$$(8) \qquad P^{-J} > C_1/c_0.$$

For P large, this holds if $J < 0$. This is true by (5_2) if

$$(9) \qquad r(1-\nu)^k < 1, \qquad r = (3-f)n/g.$$

But r increases with f since

$$dr/df = (n/g^2)(3n - 7/2 + 3\nu + z) > 0.$$

Since (9_1) is equivalent to

$$(10) \qquad k > \log r/\{-\log(1-\nu)\},$$

k will be a minimum if r and hence f is. But $f \geqq \frac{1}{2}\nu$. Hence we employ the minimum f, viz.,

$$(11) \qquad f = \tfrac{1}{2}\nu.$$

The resulting value of r in (9) is

$$(12) \qquad r = n^2(6n-1)/(n-d), \qquad d = 1 + 2n^2 z.$$

For a small z, say $\nu^3/24$, let k_0 be the least integer k satisfying (10). Then all sufficiently large integers are sums of $4n - 2 + 3k_0$ n-th powers.

3. Results which are better for universal theorems are obtained by taking [2] $k = 2k_0$, whence k exceeds the double of the second member of (10). By the definition (5) of σ and (10), $\log\sigma$ is just $< \log n - 2\log r$. Hence

$$\sigma < \frac{(n-d)^2}{n^3(6n-1)^2}, \qquad 3 - f = \frac{6n-1}{2n}, \qquad \tfrac{1}{2}\sigma(3-f) < \frac{(n-d)^2}{4n^4(6n-1)}.$$

Take $z = \nu^3/24$. Then $1 < d < 2$. Thus

$$(12d-1)n > 11n > 24 > 6d^2, \qquad (n-d)^2/(6n-1) < n/6,$$

$$-J = -\frac{1}{2}\sigma(3-f) - \frac{1}{2}z + \frac{1}{4}(\nu - \nu^2) > \frac{-1}{24n^3} - \frac{1}{48n^3} + \frac{1}{4}\left(\frac{1}{n} - \frac{1}{n^2}\right) \geqq \frac{2}{9n},$$

if $(2n-9)^2 \geqq 90$ and hence if $n \geqq 10$. Thus (8) holds if

$$(13) \qquad \log_{10} P \geqq 2n \log_e C_1/c_9, \qquad n \geqq 9,$$

since it is readily verified if $n = 9$.

4. *Evaluation of* c_9. Let S denote the singular series and

$$n \geqq 4, \qquad s \geqq 4n, \qquad b = (1 + n^s)^{2/(s-5)}.$$

Then [3] $S > b_4$, where, if the products range over the primes p indicated,

---

[2] We discard (119) of *A*, whose object was to make σ very small, a property first used in § 18, which is to be omitted now. We may omit the condition $s \geqq 70$.

[3] Landau, *Vorlesungungen über Zahlentheorie*, vol. 1, p. 303; James, *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 408, 434-435.

$$b_4 = \prod_{p \leqq b} b(p) \cdot \prod_{p > b} (1 - p^{-3/2}), \qquad b(p) > \begin{cases} p^{-n(2n-1)}, & p > 2, \\ 2^{-n(4n-1)}, & p = 2. \end{cases}$$

The second factor exceeds the like product over all primes. Hence its reciprocal is

$$< \sum_{x=1}^{\infty} x^{-3/2} \leqq 1 + \int_1^{\infty} x^{-3/2} dx = 3.$$

The reciprocal R of the first factor is

$$R < 2^{n(4n-1)} \prod_{p \leqq b} p^{n(2n-1)} \cdot 2^{-n(2n-1)}.$$

$$\log R < 2n^2 \log 2 + n(2n-1)\vartheta(b), \; \vartheta(b) < (6/5)ab + 3\log^2 b + 8 \log b + 5,$$

where[4] $\vartheta(b)$ is the sum of the natural logarithms of all primes $\leqq b$, and $a = .92129 \cdots$. Let $n \geqq 9$, $s = 36$. We increase $2n-1$ to $2n$, which more than compensates for dropping term 1 in b. Then

$$\log n \leqq n/9M, \qquad \log b \leqq 72/31 \cdot n/9M < .6n, \qquad M = .4343.$$

Thus $\log R$ will be $< n^5 - 2n^2$ if, after cancelling n^2,

$$1.3863 + 2\{(6/5)an^{72/31} + 1.08\, n^2 + 4.8\, n + 5\} < n^3 - 2.$$

Multiply the combined constant term by $1 \leqq n/9$ and likewise for the resulting terms in n and n^2. It suffices to verify

$$2.4\, an^{72/31} < .62312\, n^3, \qquad \log 2.4\, a = .3446 < .4410 = \log(.62312 \cdot 9^{21/31}).$$

When s increases, the exponent of n in b decreases, whence $\log R$ decreases. Hence, if $n \geqq 9$, $s \geqq 4n$, then $\log R < n^5 - 2n^2$. By A, $c_9 = \nu(1/3)^n b_4$. Hence

$$-\log c_9 < A + n^5 - 2n^2, \qquad A = \log n + n \log 3 + \log 3 < .18184\, n^2.$$

5. *Evaluation of* C_1. Restrict attention to the classic Waring problem. Then $\kappa = \kappa_1 = 1$, $m = 3^{n-1}$, $a = 1$. We increase C_1 by suppressing terms $-\sigma$. Thus

$$C_1 < 12(8n \cdot 3^{n-1})^{1/2} 2^{(n-1)/4} n^{3(n-1)/2} / (3/2)^{(1-\nu)/2}.$$

We increase C_1 by suppressing the denominator. Thus

$$\log_{10} C_1 < 1.53073 + (1/4)(n-1)1.25527 + ((3/2)n - 1)\log n.$$

---

[4] Landau, *Handbuch . . . Verteilung der Primzahlen*, vol. 1, p. 91.

Take $n \geqq 9$. Then $\log n \leqq n/9$. Multiply the combined constant term by $1 \leqq n/9$ and likewise for the combined term in n. Hence

$$\log_{10} C_1 < .20420\, n^2, \qquad \log_e C_1 < (1/2)n^2, \qquad n \geqq 9.$$

6. *The constant N.* Hence $\log_e C_1/c_9 < n^5$ if $n \geqq 9$. Thus (13) holds if $\log_{10} P \geqq 2n^6$. To verify that (2) is $\leqq P^z$, it suffices to use our least B. Then (2) is $< 3n^6 < P^z = 10^{n^3/12}$. Finally, P is the greatest integer $\leqq (1/3)N_0^{\nu}$. Thus (13) holds if $\log_{10} N_0 \geqq 2n^7$.

In *A*, (120), $\log_{10} c_{15} < (1/2)n^3$, $c_1 = (1/2)c_{15}{}^8$, $\log c_1 < 2n^4$. Hence in (85) the term $c_1 P^{-1}$ is insignificant. Likewise for $c_3 P^{-g}$ in (87).

THEOREM 1. *Let k_0 be the least integer exceeding* (10) *with r as in* (12). *Take $k = 2k_0$, $\log_{10} N = 2n^7$. If $n \geqq 9$, every integer $\geqq N$ is a sum of $4n - 2 + 3k$ integral n-th powers $\geqq 0$.*

PART II. THE IDEAL WARING THEOREM.

7. Write $3^n = 2^n q + r$, $0 \leqq r < 2^n$. Then $q = [(3/2)^n]$.

LEMMA 1. *All positive integers $\leqq 2^n q$ are sums of $I = 2^n + q - 2$ integral n-th powers $\geqq 0$. But $2^n q - 1$ is not a sum of $I - 1$ n-th powers.*

Let $B \leqq 2^n q - 1$. Then $B = 2^n x + y$, $0 \leqq y \leqq 2^n - 1$, and $x < q$. Thus B is a sum of $x + y \leqq 2^n - 1 + q - 1 = I$ powers.

LEMMA 2. *All integers in the interval $(2^n q, 2^n q + 2^n)$ are sums of E n-th powers, where $E = max(q + r - 1, 2^n - r)$.*

First, $2^n q, \cdots, 2^n q + r - 1$ are sums of $q + r - 1$ powers. Second, the integers $2^n q + r = 3^n, \cdots, 2^n q + 2^n - 1 = 3^n - r + 2^n - 1$ are sums of $2^n - r$ powers. Third, $2^n q + 2^n$ is a sum of $q + 1$ powers. But if $r = 1$, n would be divisible by 2^{n-2} since 3 belongs to the exponent 2^{n-2} modulo 2^n, whence $n \leqq 4$.

LEMMA 3. *Let L, p, n, z be positive integers. If all integers in the interval $(L, L + p^n z)$ are sums of m integral n-th powers $\geqq 0$, then all in the interval $(L, L + p^n z + p^n)$ are sums of $m + 1$ powers.*

Proof is needed only for integers g satisfying

$$L + p^n z < g \leqq L + p^n z + p^n, \qquad L + p^n(z-1) < g - p^n \leqq L + p^n z.$$

Since $g - p^n$ is in the first given interval, it is a sum of m powers. Hence g is a sum of $m + 1$ n-th powers.

By induction on v, Lemma 3 yields.

LEMMA 4. *Let L, p, n, z, v be positive integers. If all integers in $(L, L+p^n z)$ are sums of m integral n-th powers $\geqq 0$, then all in $(L, L+p^n(z+v))$ are sums of $m+v$ n-th powers.*

Taking $p=2$, $z=1$, and applying Lemma 2, we get

LEMMA 5. *All integers in $(2^n q, 2^n(q+y))$ are sums of $E+y-1$ n-th powers.*

Take $y=1+q$. Then $y>(3/2)^n$, $2^n y>3^n$. This proves

LEMMA 6. *All integers in $(2^n q, 2^n q+3^n)$ are sums of $E+q$ n-th powers.*

This and Lemma 4 with $p=3$, $z=1$, yield

LEMMA 7. *All in $(2^n q, 2^n q+3^n(1+v))$ are sums of $E+q+v$.*

Take $v=[(4/3)^n]$. Then $1+v>(4/3)^n$. This proves

LEMMA 8. *All in $(2^n q, 2^n q+4^n)$ are sums of $E+q+[(4/3)^n]$.*

By induction, we get

THEOREM 2. *All integers in the interval $(2^n q, 2^n q+(n+1)^n)$ are sums of*

$$E+q+\left[\left(\frac{4}{3}\right)^n\right]+\left[\left(\frac{5}{4}\right)^n\right]+\cdots+\left[\left(\frac{n+1}{n}\right)^n\right] \tag{14}$$

n-th powers, and hence are sums of $E+q+(n-2)[(4/3)^n]$ powers.

8. *Ascent.* Let $v=(1-l/L_0)/n$, $v^n L_0 \geqq 1$. If all integers between l and L_0 inclusive are sums of Q n-th powers $\geqq 0$, then [5] all between l and L_t inclusive are sums of $Q+t$ n-th powers, where

$$\log L_t=\left(\frac{n}{n-1}\right)^t(\log L_0+n\log v)-n\log v. \tag{15}$$

The interval in Theorem 2 includes (l, L_0), where $l=3^n$ and $L_0=(n+1)^n$. The above inequality reduces to $v(n+1)\geqq 1$ and hence to

$$\left(\frac{n+1}{3}\right)^{n+1}\geqq 3,$$

which holds if $n\geqq 4$. Let $n\geqq 11$. Then $v=1/n$ to seven decimal places. By the series for $\log(1+x)$, $\log L_0-n\log n$ is the product of n by

[5] Dickson, *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 711.

$$\log(1+1/n) > M(1/n - 1/2n^2), \qquad M = .4343,$$

where (as henceforth) we use logarithms to base 10.

We desire that $L_t > N$, for N in Theorem 1. This holds if

$$\left(\frac{n}{n-1}\right)^t M\left(1-\frac{1}{2n}\right) > 2n^7, \qquad n \geqq 11. \tag{16}$$

Let $n \geqq 14$. Then (16) follows from the like inequality with the denominator $2n$ replaced by 28, and hence if

$$t \log \frac{n}{n-1} > .67904 + 7 \log n.$$

The coefficient of t is the product of the modulus M by

$$\log_e\left(1+\frac{1}{n-1}\right) > \frac{1}{n-1} - \frac{1}{2(n-1)^2} > \frac{1}{n} \qquad (n > 2).$$

Hence $L_t > N$ if

$$t > M^{-1}n(.67904 + 7 \log n), \qquad n \geqq 14. \tag{17}$$

9. First, let $E = 2^n - r$. By Theorems 1 and 2, all integers $\geqq 2^n q$ are sums of

$$t + 2^n - r + q + (n-2)[(4/3)^n]$$

n-th powers, where t is the least integer satisfying (17). This number is $\leqq I = 2^n + q - 2$, if

$$r/2^n \geqq F(n), \qquad F(n) = t/2^n + (n-2)/q. \tag{18}$$

But if $E = q + r - 1$, the like conclusion holds if

$$r/2^n \leqq 1 - F(n) - q/2^n. \tag{19}$$

For $n = 14$, $t = 281$ and these limits are $F(14) = .058388$ and $.923851$. Since $F(n)$ decreases when n increases, the ideal Waring theorems holds for any $n \geqq 14$ for which $r/2^n$ lies between these decimals. Now q is the integral part and $r/2^n$ the decimal part of $(3/2)^n$. By adding the latter to one half of itself we get $(3/2)^{n+1}$. In this way we form a table of values of $(3/2)^n$ for successive n's. The above condition is seen to hold for $n = 15$-21, but not for $n = 14$.

When $n = 14$, $q = 291$, $r = 15225$, $2^n = 16384$, $t = 281$,

$$\left[\left(\frac{4}{3}\right)^{14}\right] = 56, \qquad \left[\left(\frac{5}{4}\right)^{14}\right] = 22, \qquad \left[\left(\frac{6}{5}\right)^{14}\right] < 22, \cdots.$$

Hence (14) is < 16104. Adding t, we get $16385 < I = 16673$.

Hence the ideal Waring Theorem holds for $n = 14$-21.

For $n \geqq 16$, the decimal in (17) reduces to .677034.. We get $F(21) = .00403868$, $q/2^{21} = .00237799$. Hence the ideal Waring theorem holds for any $n \geqq 21$ for which the decimal part of $(3/2)^n$ lies between .00403868 and .99358333. This is true if $21 \leqq n \leqq 162$. But for $n = 163$, the decimal part is .9955.

For $n = 163$, $F(n)$ begins with 28 zeros, while $q/2^n$ begins with 19 zeros, and hence (19) with 19 nines. Our condition holds on to $n = 180$.

THEOREM 3. *The ideal Waring theorem holds for exponents* 14-180.

10. When $n = 11$, the least integer t satisfying (16) is 184. Since $2^{11} = 2048$, $q = 86$, $r = 1019$, we have $E = 1104$. The final sum in Theorem 2 is 1397. This plus t is $1581 < I = 2132$.

11. When $n = 12$, $t = 228$. The limits (18) and (19) are .13318 and .86533. For $n = 12$ or 13, the decimal part of $(3/2)^n$ lies between these limits.

12. When $n = 10$, $\log v = \bar{2}.9999990$ and (15) gives

$$.0457575\, t \geqq 7.684117, \qquad t = 168.$$

Since $q = 57$, $r = 681$, we get $E = 737$. Then (14) is < 842. This plus t is $1010 < I = 1079$.

13. When $n = 9$, $q = 38$, $r = 227$, $E = 285$,

$$\log v = \bar{1}.045673, \qquad .0511525\, t \geqq 7.366816, \qquad t = 145,$$

$$[(4/3)^9] = 13, \qquad [(5/4)^9] = 7, \qquad (14) < 323 + 13 + 6 \times 7 = 378,$$
$$t + 378 = 523 < I = 548.$$

14. For $n = 7$ and 8, it is not possible to prove that $I = 143$ or 279 powers suffice by the preceding method, but is possible by use of results from tables.

When $n = 7$, $s = 28$, we perform the calculations in § 4 (without approximations) and get

$$\log_e 1/b_4 < 21589.453, \qquad \log_{10} c_9 = -9380.362.$$

Choose $z = .00001$. By (10) and (12), $k \geqq 38$. We take $k = 39$, which gives $4r - 2 + 3k = 143 = I$. By (5) and (6), $\sigma = .01714655$, $\log C_1 = 11.34357$. Then (7) and (8) give

$$-J = .0055, \qquad \log\log P > 6.2324, \qquad \log\log N = 7.0775.$$

But [6] 44 seventh powers suffice from $l = 14\,782\,969$ to L_0, where

$$\log L_0 = 13.8435104.$$

Here $\nu = 1/n$ to 7 decimal places. By (15) with $t = 99$, $\log\log L_t = 7.52688$. Thus $L_t > N$. Hence every integer $\geqq l$ is a sum of 143 seventh powers. The same is true for integers $\leqq l$ (*Bulletin, loc. cit.*, p. 713).

15. If $n \geqq 6$, $s \geqq 24$, we find as in §§ 4, 5 that

$$\log 1/b_4 < 2n^5 - 2n^2 + \log 3, \qquad \log C_1 < .6n^2, \qquad \log C_1/c_9 < 2n^5,$$

where the logarithms are natural.

Let $n = 8$ and take $z = .00002$. By (10) and (12), $k \geqq 46$. We take $k = 83$, for which $4n - 2 + 3k = 279 = I$. Then

$$\log \sigma = \bar{4}.089754, \qquad -J = .0271531, \qquad \log\log N = 6.92353.$$

Every integer [7] between $l = 2\,460\,866$ and $2\,851\,491$ is a sum of 81 eighth powers. By ascent, all from l to $L_0 = 2\,235\,617 \times 10^9$ are sums of 102. By (15), $L_t \geqq N$ if

$$.057992\,t + .909806 \geqq 6.92353, \qquad t \geqq 104.$$

Hence all integers $\geqq l$ are sums of I eighth powers. But (*Bulletin, loc. cit.*, p. 713), I powers suffice from 1 to far beyond l.

16. For $n = 6$, take $z = .00002$. Then $k \geqq 31$. For $k = 31$,

$$\sigma = .0210635, \qquad -J = .0039946, \qquad \log\log P > 6.2280972.$$

Adding log 6 to the last, we get $\log\log N$. In his Chicago Doctor's thesis, R. C. Shook proved that every integer between $l = 2120044$ and $L_0 = 516798 \times 10^8$ is a sum of 33 sixth powers. Apply (15), where $v = 1/n$ to 7 decimal places. Thus $L_t \geqq N$, if $t = 77$. Hence 110 powers suffice from l to N. But $4n - 2 + 3k = 115$ suffice after N, while (Shook) 86 suffice before l.

---

[6] Dickson, *Researches on Waring's Problem*, Carnegie Institution of Washington, 1935, p. 81.

[7] A. Sugar, *Chicago Master's dissertation*, 1934.

SOLUTION OF WARING'S PROBLEM.

By L. E. Dickson.

I shall prove the Waring theorem in its original sense, in contrast to an asymptotic result. For every $n > 6$ I shall evaluate $g(n)$ such that every positive integer is a sum of $g(n)$ integral n-th powers $\geqq 0$, while not all are sums of $g - 1$. Since this paper is a sequel to my preceding one, I shall continue the numbering of formulas, sections, theorems and lemmas.

17. If $[x]$ denotes the greatest integer $\leqq x$, write

$$q = [(3/2)^n], \quad f = [(4/3)^n], \quad g = [(5/4)^n], \quad s = f + 2g,$$
$$I = 2^n + q - 2.$$

Lemma 9. *If all integers in the interval $(L, L + 2^n)$ are sums of m n-th powers, then all in $(L, L + (n+1)^n)$ are sums of*

$$m + q + f + g + [(6/5)^n] + \cdots + [(n+1)^n/n^n]. \tag{20}$$

This follows from Lemma 4 as in the proof of Theorem 2.

Lemma 10. *If $n \geqq 35$ and all integers in the interval $(L, L + 2^n)$ are sums of m n-th powers, then every integer $\geqq L$ is a sum of $m + q + s - 2$ n-th powers.*

Apply[1] § 8 with $n \geqq 35$. We get (17) with the decimal replaced by .66950. Let t be the least integer satisfying (17). We shall prove that the sum of t and (20) is $< m + q + s - 1$. Cancel m, q, f, g and multiply by $(4/5)^n$. In the resulting inequality every term decreases when n increases since this is true of p^n, np^n, $n \log n \cdot p^n$, if $0 < p < 4/5$ and $n \geqq 6$. Hence it suffices to verify the initial inequality when $n = 35$, whence $t = 912$. S. S. Pillai[2] stated Lemma 10 with an undetermined limit for n, which exceeds 300 for his constant β, and is about 180 for our smaller constant N. We have

$$3^n = 2^n q + r, \qquad 1 \leqq r < 2^n. \tag{21}$$

[1] With $l = 3n$ replaced by $l = 4n$, whence in the inequality displayed below (15), 3 is replaced by 4. The new inequality holds if $n \geqq 5$. To explain these replacements, note that, in the proofs of Lemmas 12 and 14, $L < U3^n$, $U < 1 + s/4 \leqq h$, whence $L < h3^n = (3/4)4^n$. For Lemma 11, $L < 3^n$. Hence for all three lemmas, the second interval in Lemma 9 includes $(4^n, (n+1)^n)$.

[2] *Annamalai University Journal*, vol. 5 (1936), pp. 145-166. His auxiliary Lemma 13 must be altered as to the definitions of the r_i.

LEMMA 11. *If* $n \geqq 35$ *and* $s \leqq r \leqq 2^n - q - s$, *every positive integer is a sum of* I *n-th powers.*

The integers $2^n q, \cdots, 2^n q + r$ are sums of $q + r \leqq 2^n - s$ n-th powers. The integers $\geqq 2^n q + r = 3^n$ and $< 2^n(q+1)$ and hence $\leqq 3^n + 2^n - r - 1$ are sums of $2^n - r \leqq 2^n - s$ powers. Finally $2^n(q+1)$ is a sum of $q+1$ powers. Applying Lemma 10 with $L = 2^n q$, $m = 2^n s$, we see that every integer $\geqq 2^n q$ is a sum of I n-th powers. Apply also Lemma 1.

LEMMA 12. *If* $n \geqq 35$ *and* $r < s$, *every positive integer is a sum of* I *n-th powers.*

Since 3 belongs to the exponent 2^{n-2} modulo 2^n, $r = 1$ in (21) would show that n is divisible by 2^{n-2}, whence $n \leqq 4$. If $r = 3$, then $3^{n-1} = \frac{1}{3}q2^n + 1$, which was seen to be impossible unless $n - 1 \leqq 4$. Evidently r is not even. Hence $r \geqq 5$.

For $x = 0, 1, \cdots, u(3^n - 2^n q) - 1 = ur - 1$, $2^n qu + x$ is a sum of $qu + x \leqq qu + ur - 1$ n-th powers. If $2^n > ur$, there exist integers $\geqq 3^n u$ and $\leqq 2^n qu + 2^n - 1 = 3^n u + 2^n - 1 - ur$, and each is a sum of $u + 2^n - 1 - ur$ n-th powers. The latter and $qu + ur - 1$ will both be $\leqq 2^n - s$ if

$$s/(r-1) \leqq u \leqq (2^n - s)/(q+r). \tag{22}$$

We decrease r to 5 on the left and increase r to s on the right and obtain the subinterval

$$s/4 \leqq u \leqq (2^n - s)/(q+s). \tag{23}$$

Thus any u satisfying (23) will satisfy (22). The last fraction in (23) is $< 2^n/r$ since $q + s > s > r$, whence our former condition $2^n > ur$ is satisfied.

18. There will exist an integer u satisfying (23), if the difference between the limits is $\geqq 1$ and hence, if

$$2^n \geqq P, \qquad P = 2s + s^2/4 + (1 + \tfrac{1}{4}s)q.$$

But $q < (3/2)^n$, $s < 3(4/3)^n$. Hence $2^n \geqq P$, if

$$2^n > 24(4/3)^n + 9(16/9)^n + 4(3/2)^n.$$

Multiplication by $(9/16)^n$ yields the equivalent inequality

$$(9/8)^n > 9 + 24(3/4)^n + 4(27/32)^n.$$

This holds if $n = 20$ and hence if $n \geqq 20$.

Thus all integers in $(2^n qu, 2^n qu + 2^n)$ are sums of $2^n - s$ n-th powers. By Lemma 10, every integer $\geqq 2^n qu$ is a sum of I n-th powers.

19. To complete the proof of Lemma 12, it remains to prove that every positive integer $< 2^n qu$ is a sum of I n-th powers. By Lemmas 1 and 2, this is true for integers $< 3^n$. It remains to prove it for integers between 3^n and $3^n U$, where U is the least integer u satisfying (23).

LEMMA 13. *All integers in* $(3^n w, 3^n w + 2^n)$ *are sums of* M *n-th powers, where* M *is the maximum of* $A = 2^n - wr + w - 1$ *and* $B = w(q + r)$.

For, those $\leqq w3^n + 2^n - wr - 1$ are sums of A powers. The next integer is equal to $2^n(qw + 1)$. To this we have to add $1, \cdots, wr - 1$, and obtain $3^n w + 2^n - 1$ as the final sum.

When $w \leqq U - 1$, we shall prove that $B \leqq A$, viz.,

$$(24) \qquad 2^n \geqq wq + 2wr - w + 1.$$

By definition of U, $U - 1 < s/4$. Employ $W = [h]$, $h = \frac{3}{4}(4/3)^n$. Then $W > U - 1$ if

$$W + 1 > h \geqq 1 + s/4,$$

which holds since $n \geqq 35$. In (24) we replace w by h, increase r to $3(4/3)^n$ and q to $(3/2)^n$, and suppress $1 - w$. Thus (24) holds if $(9/8)^n \geqq 18$, which is true for $n \geqq 25$.

We may therefore take $M = A$ in Lemma 13. Hence by Lemma 4, all in $(3^n w, 3^n w + 2^n(v + 1))$ are sums of $A + v$ n-th powers. But $A + v \leqq I$ if $v \leqq wr - w + q - 1$. For the largest such v,

$$3^n w + 2^n(v + 1) = 3^n(w + 1) + 2^n(wr - w) - r > 3^n(w + 1).$$

But (24) implies the like inequality for a smaller w. Hence all integers in $(3^n w, 3^n(w + 1))$ are sums of I n-th powers for $w = 1, 2, \cdots, U - 1$.

LEMMA 14. *If* $n \geqq 35$ *and* $2^n - q - s < r \leqq 2^n - q - 5$, *every positive integer is a sum of* I *n-th powers.*

For $x = 0, \cdots, 3^n u - 2^n(uq + u - 1) - 1 = 2^n - u(2^n - r) - 1$, $2^n(uq + u - 1) + x$ is a sum of $C = 2^n - u(2^n - r - q - 1) - 2$ n-th powers. The integers $\geqq 3^n u$ and $\leqq 2^n(uq + u) - 1 = 3^n u + u(2^n - r) - 1$ are sums of $D = u + u(2^n - r) - 1$ n-th powers. Then C and D are both $\leqq 2^n - s$ if

$$(25) \qquad s/(2^n - r - q - 1) \leqq u \leqq (2^n - s)/(2^n - r + 1).$$

We increase r to $2^n - q - 5$ on the left and decrease r to $2^n - q - s + 1$ on the right and obtain (23) as a subinterval.

Note that the greatest x is positive if $u \leqq (2^n - 1)/(2^n - r)$, which follows from (25). The discussion in § 18 applies also here and shows that every integer $\geqq 3^n u$ is a sum of I n-th powers.

20. We shall prove that the latter is true also of the integers in $(3^n w, 3^n(w+1))$ for $w = 1, \cdots, U-1$, where U is the least integer u satisfying (23). We shall later restrict n so that

$$2^n \geqq wq + (w+1)s. \tag{26}$$

Then $2^n > wq + ws$, whence

$$wr > w(2^n - q - s) > (w-1)2^n, \qquad w(2^n - r) < 2^n.$$

As explained in various recent papers, my method to obtain the minimum decompositions of integers in $J = (3^n w, 3^n w + 2^n)$ is here based on the equations

$$\begin{array}{lll} & \text{Diff.} & \text{Wt.} \\ j\{-3^n + (q+1)2^n\} = j(2^n - r) & 2^n - r & jq \quad (j = 1, \cdots, w-1). \\ w\{-3^n + (q+1)2^n\} = w(2^n - r) & 2^n - w(2^n - r) & wq \end{array}$$

To Diff. + Wt. -1 we add w (since $w3^n$ must be added to get decompositions). For $j \leqq w-1$, the largest sum is $F = 2^n - r + (w-1)(q+1)$. For our final equation, the sum is $G = 2^n - w(2^n - r) + wq + w - 1$. Then[3] $G \geqq F$ if $(w+1)r \geqq 2^n w - q$. By an hypothesis in Lemma 14, this follows from (26). This proves that every integer in the interval J is a sum of G n-th powers. Hence all in $(3^n w, 3^n w + (y+1)2^n)$ are sums of $G + y$ powers. Use the largest y for which $G + y \leqq I$. The interval now ends with

$$3^n w + 2^n\{w(2^n - r) + (1-w)q - w\} \geqq (w+1)3^n$$

if the number in brackets is $\geqq 1 + q$ and hence is $\geqq (3/2)^n$. The condition is $w(2^n - r - q - 1) > 0$ and is satisfied.

For W below (24), $W > U - 1 \geqq w$. Hence (26) follows from the like inequality with w replaced by W, and q increased to $(3/2)^n$, and s increased to $(4/3)^n + 2(5/4)^n$. Division by $2^n/4$ yields

$$1 \geqq 4(2/3)^n + 8(5/8)^n + 3(8/9)^n + 6(5/6)^n.$$

The second member evidently decreases when n increases. But the inequality holds if $n = 23$ and hence if $n \geqq 23$.

This proof of Lemma 14 is readily extended to $r = 2^n - q - i$, $i = 3$ or 4, by employing (25) instead of (23), and using $s < (4/3)^n + 2(5/4)^n$ in § 18.

[3] Expressed untechnically, the minimum decompositions are free of terms 3^n.

Lemmas 11-14 yield

THEOREM 4. *If $n \geqq 35$ and $r \leqq 2^n - q - 3$, every positive integer is a sum of I n-th powers.*

This inequality has been verified for $4 \leqq n \leqq 180$.

21. Let $r \geqq 2^n - q$. If $r = 2^n - 1$, (21) gives $3^{2n} \equiv 1 \pmod{2^n}$. Since 3 belongs to the exponent 2^{n-2}, $2n$ is divisible by 2^{n-2}, but is less than it if $n \geqq 6$. Next, if $r = 2^n - 3$, (21) gives $3^{n-1} = \frac{1}{3}(q+1)2^n - 1$, whence $n - 1 \leqq 5$. Since r is odd, we conclude that $r \leqq 2^n - 5$. Thus

$$(27) \qquad 5 \leqq R \leqq q, \qquad R = 2^n - r.$$

Since many integers $< 4^n$ require more than I summands (§ 22), we shall make our ascents from an interval containing 4^n. Let

$$(28) \qquad 4^n = 3^n f + 2^n h + j, \qquad 0 \leqq 2^n h + j < 3^n, \qquad 0 \leqq j < 2^n.$$

Then $f = [(4/3)^n]$. Fortunately we can determine h and j in terms of q, r, f. In (28) replace 3^n by its value (21). Then $j + fr \equiv 0$, $j - fR \equiv 0 \pmod{2^n}$. But $fR \leqq fq < (4/3 \cdot 3/2)^n$. Hence gR and j are numbers x for which $0 \leqq x < 2^n$. Thus $|j - fR| < 2^n$. The multiple $j - fR$ of 2^n is therefore zero. Eliminating 3^n, j, r from (28) by use of (21) and (27), we get $4^n = 2^n(fq + h + f)$. Hence

$$(29) \qquad j = fR, \qquad h = 2^n - fq - f.$$

For these values, the first inequality (28) is satisfied.

If $E = \max(A = 2^n - f(q - R) - 1, 2^n - fR)$, every integer in the interval $(3^n f + 2^n h, 3^n f + 2^n(h+1))$ is a sum of E n-th powers, as shown by

$$\begin{array}{ll} 3^n f + 2^n h + x, & (x = 0, 1, \cdots, fR - 1), \\ 4^n + y, & (y = 0, 1, \cdots, 2^n - fR - 1). \end{array}$$

By Theorem 2, all in $(4^n, (n+1)^n)$ are sums of (14) n-th powers.

Let $n \geqq 35$. Let t be the least integer satisfying (17) with the decimal replaced by .66950.

First, let $E = 2^n - fR$. Then all integers $\geqq 4^n$ are sums of

$$t + 2^n - fR + q + f + K, \qquad K = [(5/4)^n] + \cdots + [(n+1)^n/n^n].$$

Since $K < (n-3)(5/4)^n$, this sum is $< I$ if

$$R - 1 \geqq (t+2)(3/4)^n + (n-3)(15/16)^n.$$

The latter decreases when n increases. For $n = 35$ it becomes $R \geqq 4.3816$ since $t = 912$. Hence it holds when $n \geqq 35$.

Second, let $E = A$. Then all integers $\geqq 4^n$ are sums of $t + A + q + f + K$ powers. This will be $\leqq I$, if $f(q - R - 1) \geqq t + 1 + K$. If $n = 35$, this holds if $q - R - 1 \geqq .21$. Hence if $n \geqq 35$, it holds if $R \leqq q - 2$.

THEOREM 5. *If $n \geqq 35$ and $r \geqq 2^n - q + 2$, every integer $\geqq 4^n$ is a sum of I n-th powers.*

22. *Minimum decomposition of integers between 3^n and 4^n when $r \geqq 2^n - q$.* My general theory (§ 20) here has the peculiarity that there is no effective equation. For, such an equation is a multiple of $(q + 1)2^n = 3^n + 2^n - r$. The weight $q + 1$ of the left member is $\geqq$ the weight $1 + 2^n - r$ of the right. Expressed otherwise every minimum decomposition of an integer $< 4^n$ is a normal decomposition $3^n x + 2^n y + z$ in which x, y, z are integers $\geqq 0$ and x is a maximum, so that $2^n y + z < 3^n$, $z < 2^n$. Comparison with (28) and (21) gives $x \leqq f$, $y \leqq q$. Write $S = x + y + z$.

If $y = q$, then $z < r$. But $r \leqq 2^n - 5$ (§ 20). Hence $S < I + f - 3$.

Let $x = f$. Then $y \leqq h$. But $h \leqq q$. If $y = h$, then $z < fR < 2^n$, and $S \leqq I + f$, with equality only when $h = q$, $fR = 2^n - 1$. Next, if $y \leqq h - 1$, $S \leqq f + h + 2^n - 2 \leqq I + f$, and $S = I + f$ only when $y = h - 1$, $h = q$.

Finally, if $x \leqq f - 1$ and $y \leqq q - 1$, then $S \leqq I + f - 1$.

Hence if $r \geqq 2^n - q$, every integer $\leqq 4^n$ is a sum of $I + f$ n-th powers. The integer $3^n(f - 1) + 2^n(q - 1) + 2^n - 1$ is a sum of $I + f - 1$, but not fewer, n-th powers. Only in the special case $h = q$ does there exist an integer requiring $I + f$ powers.

These results and Theorem 5 yield

THEOREM 6. *If $n \geqq 35$ and $r \geqq 2^n - q + 2$, then $g(n) = I + f$ or $I + f - 1$ according as $2^n = fq + f + q$ or $2^n < fq + f + q$.*

23. To prove Theorem 6 also when $r = 2^n - q + 1$, whence $R = q - 1$, we permit $I + f - 1$ summands instead of I in the second case $E = A$ of § 21. The condition is now $f(q - R) \geqq t + 2 + K$, which holds if $q - R \geqq .21$ and hence if $R \leqq q - 1$.

For $r = 2^n - q$ (or $R = q$) a like proof applies if we permit $I + f(1 + d)$ summands, where $d = .21$ if $n \geqq 35$, $d = .001614$ if $n \geqq 180$. But if we permit only $I + f - 1$ summands, we must start our ascents from an interval beyond 4^n.

24. There remain only the cases $r = 2^n - q - i$, $i = 1, 2$.

If $r = 2^n - q - 1$, (21) gives $3^n = (q + 1)(2^n - 1)$. Hence $2^n - 1 = 3^t$, $t \leqq n$, $3^{2t} \equiv 1 \pmod{2^n}$. Since 3 belongs to the exponent 2^{n-2}, the latter divides $2t$. But $2t \leqq 2n < 2^{n-2}$ if $n \geqq 6$. Hence $t = 0$, $n = 1$.

THE UNIVERSITY OF CHICAGO.

ON THE NUMBER OF REPRESENTATIONS OF AN INTEGER AS A SUM OF $2h$ SQUARES.[1]

By R. D. James.

1. *Introduction.* Let $N(2^\alpha m, 2h)$ denote the number of representations of an integer $2^\alpha m$, $\alpha \geqq 0$, m an odd[2] integer > 0, as a sum of $2h$ squares. Both the arrangement of the squares and the signs of their square roots are relevant in counting the representations. Let $M(2^\alpha m, 2h, b)$ denote the number of representations of $2^\alpha m$ as a sum of $2h$ squares, exactly b of which are odd and occupy the first b places in the representations. Further, let

$$\xi_r(m) = \sum_{d|m} (-1|d)\, d^r, \qquad \sigma_r(m) = \sum_{d|m} d^r,$$

where $(-1|d) = (-1)^{(d-1)/2}$ is the Jacobi symbol.

In this paper we prove the following results.

Theorem 1. *For $\alpha \geqq 3$, $h \geqq 2$, we have*

$$(1.11)\quad M(2^\alpha m, 2h, 4s) = 2^{2s} \sum_{s/2 \leqq v \leqq (h-s)/2} \binom{h-2s}{2v-s} M(2^{\alpha-1}m, 2h, 4v).$$

For $h \geqq 2$ we have

$$(1.12)\quad M(4m, 2h, 4s) = 2^{2s} \sum_{(s-1)/2 \leqq v \leqq (h-s-1)/2} \binom{h-2s}{2v-s+1} M(2m, 2h, 4v+2).$$

For $h \geqq 2$ and $m \equiv 1 \pmod 4$ we have

$$(1.13)\quad M(2m, 2h, 4s+2) = 2^{2s+1} \sum_{s/2 \leqq v \leqq (h-s-1)/2} \binom{h-2s-1}{2v-s} M(m, 2h, 4v+1).$$

For $h \geqq 2$ and $m \equiv 3 \pmod 4$ we have

$$(1.14)\quad M(2m, 2h, 4s+2) = 2^{2s+1} \sum_{(s-1)/2 \leqq v \leqq (h-s-2)/2} \binom{h-2s-1}{2v-s+1} M(m, 2h, 4v +$$

[1] Presented to the American Mathematical Society, November 30, 1935 and December 31, 1935 under the titles "The number of representations of an integer as a sum of six, ten, or fourteen squares" and "The number of representations of an integer as a sum of twelve, sixteen, or twenty squares."

[2] Throughout this paper m will always denote an odd integer.

THEOREM 2. *Let α and β be integers such that $\alpha \geqq \beta + 2 \geqq 2$. Then, given the integers $A_s(h)$, $1 \leqq s \leqq (h-1)/2$, there exist integers $A_v(h,\beta)$, $1 \leqq v \leqq (h-1)/2$, $0 \leqq \beta \leqq \alpha - 2$ such that*

$$(1.21) \qquad \sum_{1 \leqq s \leqq (h-1)/2} A_s(h) M(2^\alpha m, 2h, 4s) = \sum_{1 \leqq v \leqq (h-1)/2} A_v(h,\beta) M(2^{\alpha-\beta} m, 2h, 4v).$$

The integers $A_v(h,\beta)$ are defined by the formulas

$$A_v(h,0) = A_v(h),$$

$$(1.22) \qquad A_v(h,\beta) = \sum_{1 \leqq s \leqq \min(2v, h-2v)} 2^{2s} \binom{h-2s}{2v-s} A_s(h, \beta - 1).$$

Moreover they satisfy the equation

$$(1.23) \qquad \sum_{1 \leqq v \leqq (h-1)/2} A_v(h,\beta) = 2^{(h-1)\beta} \sum_{1 \leqq s \leqq (h-1)/2} A_s(h).$$

THEOREM 3. *For $\alpha \geqq 0$, $m > 0$, $t \geqq 1$ there exist integers $\beta(t)$, $\gamma_r(t)$, $0 \leqq r \leqq t-1$, such that*

$$(1.31) \qquad \begin{aligned} &\beta(t) N(2^\alpha m, 4t+2) - 4[(-1)^t(-1|m) 2^{2t(\alpha+1)} + 1]\xi_{2t}(m) \\ &\quad = \sum_{r=0}^{t-1} \gamma_r(t) M(2^{\alpha+2} m, 4t+2, 4t-4r), \end{aligned}$$

with

$$(1.32) \qquad \sum_{r=0}^{t-1} \gamma_r(t) = 0.$$

THEOREM 4. *For $\alpha \geqq 0$, $m > 0$, $t \geqq 1$ there exist integers $\lambda(t)$, $\mu_r(t)$, $0 \leqq r \leqq t-1$, such that*

$$(1.41) \qquad \begin{aligned} &\lambda(t) N(2^\alpha m, 4t+4) \\ &\quad - 2^{2t+3}\left[2^{(2t+1)\alpha} - (-1)^t \frac{2^{(2t+1)\alpha} - 2^{2t+2} + 1}{2^{2t+1} - 1}\right] \sigma_{2t+1}(m) \\ &\quad = \sum_{r=0}^{t-1} \mu_r(t) M(2^{\alpha+2} m, 4t+4, 4r+4), \end{aligned}$$

with

$$(1.42) \qquad \sum_{r=0}^{t-1} \mu_r(t) = 0.$$

Theorem 1 is proved by a straightforward examination of the relations between $M(2^\alpha m, 2h, 4s)$ and $M(2^{\alpha-1} m, 2h, 4v)$. Theorem 2 is a consequence of Theorem 1. In applications the condition (1.23) is the important part of this theorem. Theorems 3 and 4 depend on results from the theory of theta-functions. They are obtained in the usual way by equating the coefficients of two power series representing the same function. The conditions (1.32) and (1.42) are interesting results.

As examples of the application of the above general theorems to particular problems we quote the following results.

THEOREM 5. *For no $\alpha \geqq 1$ is an equation of the form*

$$N(2^{\alpha}m, 14) = b_6(\alpha)\xi_6(m)$$

valid for all odd integers m with $b_6(\alpha)$ independent of m.

THEOREM 6. *For no $\alpha \geqq 1$ is an equation of the form*

$$N(2^{\alpha}m, 4t+4) = c_{2t+1}(\alpha)\sigma_{2t+1}(m), \qquad (t = 3, 4, \text{ or } 5),$$

valid for all odd integers m with $c_{2t+1}(\alpha)$ independent of m.

Theorems 5 and 6 are extensions of results proved by E. T. Bell.[3] He showed that no equations of the form $N(m, 4t+2) = b_{2t}\xi_{2t}(m)$, $t \geqq 3$ or $N(2m, 4t+4) = c_{2t+1}\sigma_{2t+1}(m)$, $t \geqq 3$ are possible for all odd integers m with b_{2t} and c_{2t+1} independent of m. J. W. L. Glaisher[4] gave results of the type of Theorems 5 and 6 but without proof.

2. *Proof of Theorem* 1. For a, b, h, and s integers such that $0 \leqq 2s \leqq h$, $0 \leqq a \leqq 2s$, $0 \leqq b \leqq h - 2s$ let $M_{ab}(2^{\alpha}m, 2h, 4s)$ denote the number of solutions of the equation

$$2^{\alpha}m = \sum_{i=1}^{4s} x_i^2 + \sum_{j=1}^{2h-4s} y_j^2 \tag{2.11}$$

with the added restrictions

$$x_i \text{ odd}, \qquad y_j \text{ even}; \tag{2.12}$$

$$\begin{aligned} x_{2\mu-1} + x_{2\mu} &\equiv 2 \pmod 4, & (\mu &= 1, \cdots, a); \\ x_{2\mu-1} + x_{2\mu} &\equiv 0 \pmod 4, & (\mu &= a+1, \cdots, 2s); \end{aligned} \tag{2.13}$$

$$\begin{aligned} y_{2\nu-1} + y_{2\nu} &\equiv 2 \pmod 4, & (\nu &= 1, \cdots, b); \\ y_{2\nu-1} + y_{2\nu} &\equiv 0 \pmod 4, & (\nu &= b+1, \cdots, h-2s). \end{aligned} \tag{2.14}$$

We have the following result.

LEMMA 1. *For $\alpha \geqq 1$ we have*

$$M_{ab}(2^{\alpha}m, 2h, 4s) = M(2^{\alpha-1}m, 2h, 2s+2b).$$

[3] *Journal of the London Mathematical Society*, vol. 4 (1929), pp. 279-285; *Journal für die reine und angewandte Mathematik*, vol. 163 (1930), pp. 65-70.

[4] *Quarterly Journal of Mathematics*, vol. 38 (1906-7), pp. 178-237; *Proceedings of the London Mathematical Society*, Series 2, vol. 5 (1907), pp. 479-490.

Proof. Write

$$
(2.21)\quad
\begin{aligned}
X_\mu &= \tfrac{1}{2}(x_{2\mu-1} + x_{2\mu}), & (\mu &= 1,\cdots,a);\\
X_\mu &= \tfrac{1}{2}(x_{2\mu-1} - x_{2\mu}), & (\mu &= a+1,\cdots,2s);\\
X_{2s+2\nu-1} &= \tfrac{1}{2}(y_{2\nu-1} + y_{2\nu}), & (\nu &= 1,\cdots,b);\\
X_{2s+2\nu} &= \tfrac{1}{2}(y_{2\nu-1} - y_{2\nu}), & (\nu &= 1,\cdots,b);\\
Y_\mu &= \tfrac{1}{2}(x_{2\mu-1} - x_{2\mu}), & (\mu &= 1,\cdots,a);\\
Y_\mu &= \tfrac{1}{2}(x_{2\mu-1} + x_{2\mu}), & (\mu &= a+1,\cdots,2s);\\
Y_{2s+2\nu-2b-1} &= \tfrac{1}{2}(y_{2\nu-1} + y_{2\nu}), & (\nu &= b+1,\cdots,h-2s);\\
Y_{2s+2\nu-2b} &= \tfrac{1}{2}(y_{2\nu-1} - y_{2\nu}), & (\nu &= b+1,\cdots,h-2s).
\end{aligned}
$$

Using (2.12), (2.13), and (2.14) we find that the X's are odd integers and the Y's are even integers. It then follows from (2.11) and (2.21) that

$$
(2.22)\quad 2^{a-1}m = \sum_{i=1}^{2s+2b} X_i^2 + \sum_{j=1}^{2h-2s-2b} Y_j^2, \qquad X_i \text{ odd}, \; Y_j \text{ even}.
$$

Conversely write

$$
(2.23)\quad
\begin{aligned}
x_{2\mu-1} &= X_\mu + Y_\mu, & (\mu &= 1,\cdots,a);\\
x_{2\mu} &= X_\mu - Y_\mu, & (\mu &= 1,\cdots,a);\\
x_{2\mu-1} &= X_\mu + Y_\mu, & (\mu &= a+1,\cdots,2s);\\
x_{2\mu} &= Y_\mu - X_\mu, & (\mu &= a+1,\cdots,2s);\\
y_{2\nu-1} &= X_{2s+2\nu-1} + X_{2s+2\nu}, & (\nu &= 1,\cdots,b);\\
y_{2\nu} &= X_{2s+2\nu-1} - X_{2s+2\nu}, & (\nu &= 1,\cdots,b);\\
y_{2\nu-1} &= Y_{2s+2\nu-2b-1} + Y_{2s+2\nu-2b}, & (\nu &= b+1,\cdots,h-2s);\\
y_{2\nu} &= Y_{2s+2\nu-2b-1} - Y_{2s+2\nu-2b}, & (\nu &= b+1,\cdots,h-2s).
\end{aligned}
$$

Then from (2.22) we obtain (2.11) with the restrictions (2.12), (2.13), and (2.14). It is easily seen from (2.21) and (2.23) that distinct solutions of (2.11) give rise to distinct solutions of (2.22) and conversely. Hence the number of solutions of (2.11) with the restrictions (2.12), (2.13), and (2.14) is equal to the number of solutions of (2.22). Equation (2.22), however, has $M(2^{a-1}m, 2h, 2s+2b)$ solutions and the lemma is proved.

The proof of Theorem 1 now proceeds as follows. To obtain the solutions of (2.11) with only the restrictions (2.12) we must add up the number of solutions $M_{ab}(2m, 2h, 4s)$ for $0 \leqq a \leqq 2s$, and $0 \leqq b \leqq h-2s$, taking into account the fact that a different arrangement of the x's and y's constitutes a different solution. Thus each solution of (2.11) with the restrictions (2.12), (2.13), and (2.14) counts $\binom{2s}{a}\binom{h-2s}{b}$ times as a solution of (2.11) with only the restriction (2.12). Hence by Lemma 1 we have

$$
\begin{aligned}
M(2^{a}m, 2h, 4s) &= \sum_{a=0}^{2s} \sum_{b=0}^{h-2s} \binom{2s}{a} \binom{h-2s}{b} M_{ab}(2^{a}m, 2h, 4s) \\
(2.24) \qquad &= \sum_{a=0}^{2s} \binom{2s}{a} \sum_{b=0}^{h-2s} \binom{h-2s}{b} M(2^{a-1}m, 2h, 2s+2b) \\
&= 2^{2s} \sum_{b=0}^{h-2s} \binom{h-2s}{b} M(2^{a-1}m, 2h, 2s+2b).
\end{aligned}
$$

A sum of $2s + 2b$ odd squares is congruent to $2s + 2b \pmod 4$ and a sum of even squares is congruent to $0 \pmod 4$. Therefore if $\alpha \geqq 3$ we have $M(2^{a-1}m, 2h, 2s+2b) = 0$ unless $2s + 2b \equiv 0 \pmod 4$, that is, $2s + 2b = 4v$. Then from (2.24) with $2b$ replaced by $4v - 2s$ we obtain (1.11).

The remaining formulas are proved in exactly the same way. The only changes are that we must have $2s + 2b = 4v + 2$, $4v + 1$, and $4v + 3$ in (1.12), (1.13), and (1.14), respectively.

3. *Proof of Theorem* 2. From Theorem 1 we have

$$
\begin{aligned}
&\sum_{1 \leqq s \leqq (h-1)/2} A_s(h) M(2^{a}m, 2h, 4s) \\
&\quad = \sum_{1 \leqq s \leqq (h-1)/2} 2^{2s} \sum_{s/2 \leqq v \leqq (h-s)/2} \binom{h-2s}{2v-s} M(2^{a-1}m, 2h, 4v) \\
&\quad = \sum_{1 \leqq v \leqq (h-1)/2} M(2^{a-1}m, 2h, 4v) \sum_{1 \leqq s \leqq \min(2v, h-2v)} 2^{2s} \binom{h-2s}{2v-s} A_s(h) \\
&\quad = \sum_{1 \leqq v \leqq (h-1)/2} A_v(h, 1) M(2^{a-1}m, 2h, 4v),
\end{aligned}
$$

where

$$
\begin{aligned}
\sum_{1 \leqq v \leqq (h-1)/2} A_v(h, 1) &= \sum_{1 \leqq v \leqq (h-1)/2} \; \sum_{1 \leqq s \leqq \min(2v, h-2v)} 2^{2s} \binom{h-2s}{2v-s} A_s(h) \\
&= \sum_{1 \leqq s \leqq (h-1)/2} 2^{2s} A_s(h) \sum_{s/2 \leqq v \leqq (h-s)/2} \binom{h-2s}{2v-s} \\
&= 2^{h-1} \sum_{1 \leqq s \leqq (h-1)/2} A_s(h).
\end{aligned}
$$

This proves (1.21) and (1.23) for the case $\beta = 1$.

We complete the proof by induction. Assume that we have

$$
\begin{aligned}
&\sum_{1 \leqq s \leqq (h-1)/2} A_s(h) M(2^{a}m, 2h, 4s) \\
&\qquad = \sum_{1 \leqq v \leqq (h-1)/2} A_v(h, \beta - 1) M(2^{a-\beta+1}m, 2h, 4v)
\end{aligned}
$$

with

$$
\sum_{1 \leqq v \leqq (h-1)/2} A_v(h, \beta - 1) = 2^{(h-1)(\beta-1)} \sum_{1 \leqq s \leqq (h-1)/2} A_s(h).
$$

Then, using Theorem 1 with α replaced by $\alpha - \beta + 1$ we obtain (1.21), (1.22), and (1.23).

4. *Proof of Theorems* 3 *and* 4. The methods of proof of Theorems 3 and

4 are similar and we give the details only for Theorem 3. The elliptic function $\mathrm{dn}(u, k)$ has the expansion [5]

$$\mathrm{dn}(u,k) = 1 + \sum_{t=1}^{\infty} \frac{(-1)^t u^{2t}}{(2t)!} \left[\sum_{r=0}^{t-1} d_r(t) k^{2t-2r} \right], \tag{4.11}$$

where the $d_r(t)$ are integers > 0. In the usual notation for the theta-functions we have

$$\begin{aligned}
\vartheta_0(x) &= \vartheta_0(x,q) = 1 + 2\sum_{n=1}^{\infty} (-1)^n q^{n^2} \cos 2nx, \\
\vartheta_2(x) &= \vartheta_2(x,q) = 2\sum_{n=0}^{\infty} q^{(2n+1)^2/4} \cos (2n+1)x, \\
\vartheta_3(x) &= \vartheta_3(x,q) = 1 + 2\sum_{n=1}^{\infty} q^{n^2} \cos 2nx, \\
\vartheta_0 &= \vartheta_0(0), \qquad \vartheta_2 = \vartheta_2(0) = 2\sum_{n=0}^{\infty} q^{(2n+1)^2/4}, \\
\vartheta_3 &= \vartheta_3(0) = 1 + 2\sum_{n=1}^{\infty} q^{n^2}.
\end{aligned} \tag{4.12}$$

The function $\mathrm{dn}(u, k)$ is related to the theta quotients by the formulas

$$\frac{\vartheta_3(x)}{\vartheta_0(x)} = \frac{\vartheta_3}{\vartheta_0} \mathrm{dn}(\vartheta_3{}^2 x, \vartheta_2{}^2/\vartheta_3{}^2), \tag{4.13}$$

$$\frac{\vartheta_3(x)}{\vartheta_2(x)} = \frac{\vartheta_3}{\vartheta_2} \mathrm{dn}(i\vartheta_3{}^2 x, \vartheta_0{}^2/\vartheta_3{}^2). \tag{4.14}$$

It follows from (4.11), (4.13), and (4.14) that

$$\frac{\vartheta_3(x)}{\vartheta_0(x)} = \frac{\vartheta_3}{\vartheta_0} \left\{ 1 + \sum_{t=1}^{\infty} \frac{(-1)^t x^{2t}}{(2t)!} \left[\sum_{r=0}^{t-1} d_r(t) \vartheta_3{}^{4r} \vartheta_2{}^{4t-4r} \right] \right\}, \tag{4.21}$$

$$\frac{\vartheta_3(x)}{\vartheta_2(x)} = \frac{\vartheta_3}{\vartheta_2} \left\{ 1 + \sum_{t=1}^{\infty} \frac{x^{2t}}{(2t)!} \left[\sum_{r=0}^{t-1} d_r(t) \vartheta_3{}^{4r} \vartheta_0{}^{4t-4r} \right] \right\}. \tag{4.22}$$

We can, however, obtain these expansions in another way. It is known that [6]

$$\frac{\vartheta_3(x)}{\vartheta_0(x)} = \frac{1}{\vartheta_0\vartheta_3} \left\{ 1 + 4\sum_{n=1}^{\infty} q^n \left[\sum (-1|\tau) \cos (2nx/\tau) \right] \right\}, \tag{4.31}$$

$$\frac{\vartheta_3(x)}{\vartheta_2(x)} = \frac{1}{\vartheta_2\vartheta_3} \left\{ \sec x + 4\sum_{n=1}^{\infty} q^n \left[\sum (-1|\tau) \cos \tau x \right] \right\}, \tag{4.32}$$

where the summation for τ is over all *odd* divisors of n. By expanding the cosines in (4.31) and (4.32) and then comparing the result with (4.21) and (4.22) we obtain

---

[5] See, for example, Tannery and Molk, *Éléments de la Théorie des Fonctions elliptiques*, Paris (1902), Vol. 4, p. 92.

[6] E. T. Bell, *Messenger of Mathematics*, vol. 49 (1919-20), pp. 78-84.

$$(4.33)\quad \vartheta_3^2 \sum_{r=0}^{t-1} d_r(t)\vartheta_3^{4r}\vartheta_2^{4t-4r} = 2^{2t+2}\sum_{n=1}^{\infty} q^n \Big[\sum (-1|\tau)(n/\tau)^{2t}\Big];$$

$$(4.34)\quad \vartheta_3^2 \sum_{r=0}^{t-1} d_r(t)\vartheta_3^{4r}\vartheta_0^{4t-4r} = K_{2t} + 4(-1)^t \sum_{n=1}^{\infty} q^n \Big[\sum (-1|\tau)\tau^{2t}\Big].$$

We use these formulas to prove the following result.

LEMMA 2. *For $\alpha \geqq 0$, $m > 0$, $t \geqq 1$ we have*

$$(4.41)\quad \sum_{r=0}^{t-1} d_r(t) M(2^{\alpha+2}m, 4t+2, 4t-4r) = 2^{2t(\alpha+1)+2}(-1|m)\xi_{2t}(m),$$

$$(4.42)\quad \sum_{r=0}^{t} h_r(t) M(2^{\alpha+2}m, 4t+2, 4t-4r) = 4\xi_{2t}(m),$$

where

$$(4.43)\qquad h_r(t) = (-1)^r \sum_{s=0}^{\min(r,t-1)} \binom{t-s}{r-s} d_s(t).$$

Proof. From (4.12) it follows that

$$(4.44)\quad \begin{aligned} \vartheta_3^{4r+2}\vartheta_2^{4t-4r} &= \Big(1+2\sum_{n=1}^{\infty} q^{n^2}\Big)^{4r+2}\Big(2\sum_{n=0}^{\infty} q^{(2n+1)^2/4}\Big)^{4t-4r} \\ &= \sum_{n=0}^{\infty} q^{n/4} M(n, 4t+2, 4t-4r). \end{aligned}$$

Since a sum of $4t-4r$ odd squares is congruent to $4t-4r \equiv 0 \pmod 4$ and a sum of even squares is congruent to $0 \pmod 4$, we have $M(n, 4t+2, 4t-4r) = 0$ unless n is divisible by 4. Hence (4.44) becomes

$$(4.45)\qquad \vartheta_3^{4r+2}\vartheta_2^{4t-4r} = \sum_{n=0}^{\infty} q^n M(4n, 4t+2, 4t-4r).$$

Comparing (4.45) and (4.33) for $n = 2^\alpha m$ we obtain

$$(4.46)\quad \sum_{r=0}^{t-1} d_r(t) M(2^{\alpha+2}m, 4t+2, 4t-4r) = 2^{2t+2}\sum (-1|\tau)(2^\alpha m/\tau)^{2t},$$

where the summation for τ is over all *odd* divisors of $2^\alpha m$, that is, over *all* divisors of m. To complete the proof of (4.41) it remains to show that the right sides of (4.41) and (4.46) are equal. We have [7]

$$\begin{aligned} \sum_{\tau|m} (-1|\tau)(m/\tau)^{2t} &= (-1|m)\sum_{\tau|m}(-1|(m/\tau))(m/\tau)^{2t} \\ &= (-1|m)\sum_{d|m}(-1|d)d^{2t} = (-1|m)\xi_{2t}(m). \end{aligned}$$

To prove (4.42) we first employ the well known relation $\vartheta_0^4 = \vartheta_3^4 - \vartheta_2^4$.

[7] It is a property of the Jacobi symbol that $(-1 \mid PQ) = (-1 \mid P)\ (-1 \mid Q)$.

and substitute for ϑ_0^{4t-4r} in (4.34). On expanding and collecting terms we obtain

$$\sum_{r=0}^{t} h_r(t)\vartheta_3^{4r+2}\vartheta_2^{4t-4r} = 4\sum_{n=1}^{\infty} q^n \left[\sum (-1|\tau)\tau^{2t} \right],$$

where the $h_r(t)$ are given by (4.43). The proof then proceeds in a similar way to that for (4.41).

Theorem 3 now follows immediately. In Lemma 2 multiply (4.41) by $(-1)^t$ and add the result to (4.42). This gives

$$\sum_{r=0}^{t-1} ((-1)^t d_r(t) + h_r(t))M(2^{\alpha+2}m, 4t+2, 4t-4r) \\ + h_t(t)M(2^{\alpha+2}m, 4t+2, 0) = 4[(-1)^t(-1|m)2^{2t(\alpha+1)} + 1]\xi_{2t}(m).$$

If we write $\beta(t) = h_t(t)$, $\gamma_r(t) = -\{(-1)^t d_r(t) + h_r(t)\}$ we obtain (1.31) of Theorem 3 since $M(2^{\alpha+2}m, 4t+2, 0) = N(2^\alpha m, 4t+2)$. It remains to prove (1.32). We have

$$\begin{aligned} \sum_{r=0}^{t-1} h_r(t) &= \sum_{r=0}^{t-1} \sum_{s=0}^{\min(r,t-1)} (-1)^r \binom{t-s}{r-s} d_s(t) \\ &= \sum_{s=0}^{t-1} d_s(t) \sum_{r=s}^{t-1} (-1)^r \binom{t-s}{r-s} \\ &= \sum_{s=0}^{t-1} d_s(t) \sum_{r=0}^{t-s-1} (-1)^{s-r} \binom{t-s}{r} \\ &= \sum_{s=0}^{t-1} d_s(t)(-1)^s\{(1-1)^{t-s} - (-1)^{t-s}\} \\ &= -\sum_{s=0}^{t-1} (-1)^s d_s(t). \end{aligned}$$

Hence

$$\sum_{r=0}^{t-1} \gamma_r(t) = -\sum_{r=0}^{t-1} (-1)^t d_r(t) - \sum_{r=0}^{t-1} h_r(t) = 0.$$

As remarked above Theorem 4 is proved in a similar way. The difference is that we use the functions $\vartheta'_0(x)/\vartheta_0(x)$ and $\vartheta'_2(x)/\vartheta_2(x)$ instead of $\vartheta_3(x)/\vartheta_0(x)$ and $\vartheta_3(x)/\vartheta_2(x)$.

5. *Proof of Theorem* 5. From Theorem 3 with $t = 3$ we have

$$\begin{aligned} \beta(3)N(2^\alpha m, 14) &+ 4[(-1|m)2^{6\alpha+6} - 1]\xi_6(m) \\ &= \sum_{r=0}^{2} \gamma_r(3)M(2^{\alpha+2}m, 14, 12-4r), \end{aligned} \tag{5.1}$$

with $\sum_{r=0}^{2} \gamma_r(3) = 0$. Then, from Theorem 2 with $h = 7$, α replaced by $\alpha+2$, $\beta = \alpha$, $A_s(7) = \gamma_{3-s}(3)$, the right side of (5.1) becomes

$$\sum_{v=1}^{3} A_v(7,\alpha) M(4m, 14, 4v), \tag{5.2}$$

where

$$\sum_{v=1}^{3} A_v(7,\alpha) = 2^{6\alpha} \sum_{r=0}^{2} \gamma_r(3) = 0.$$

By applying first (1.12) and then (1.13) or (1.14) of Theorem 1 the expression (5.2) further reduces to

$$\begin{aligned} 8[\{255A_1(7,\alpha) + 256A_2(7,\alpha) + 256A_3(7,\alpha)\}\{M(m,14,5) + M(m,14,9)\} \\ + A_1(7,\alpha)\{M(m,14,1) + M(m,14,13)\}] \end{aligned}$$

or

$$\begin{aligned} 16[\{23A_1(7,\alpha) + 24A_2(7,\alpha) + 32A_3(7,\alpha)\}\{M(m,14,3) + M(m,14,11)\} \\ + \{210A_1(7,\alpha) + 208A_2(7,\alpha) + 192A_3(7,\alpha)\}M(m,14,7)] \end{aligned}$$

according as $m \equiv 1 \pmod 4$ or $m \equiv 3 \pmod 4$. Finally, with the condition $\sum_{v=1}^{3} A_v(7,\alpha) = 0$, these expressions become

$$\begin{aligned} 8A_1(7,\alpha)[M(m,14,1) - M(m,14,5) - M(m,14,9) + M(m,14,13)], \\ -[144A_1(7,\alpha) + 128A_2(7,\alpha)][M(m,14,3) - 2M(m,14,7) + M(m,14,11)] \end{aligned}$$

in the respective cases. Hence from (5.1) we obtain

$$\begin{aligned} &\beta(3)N(2^\alpha m, 14) + 4[(-1|m)2^{6\alpha+6} - 1]\xi_6(m) \\ &= [g(\alpha)/g(0)][\beta(3)N(m,14) + 4[(-1|m)2^6 - 1]\xi_6(m)], \end{aligned} \tag{5.4}$$

where

$$g(\alpha) = \begin{cases} 8A_1(7,\alpha), & m \equiv 1 \pmod 4, \\ -[144A_1(7,\alpha) + 128A_2(7,\alpha)], & m \equiv 3 \pmod 4. \end{cases}$$

Since it is known that $N(m,14) = b_6\xi_6(m)$ is impossible[8] with b_6 independent of m it then follows from (5.4) that there is no α for which $N(2^\alpha m, 14) = b_6(\alpha)\xi_6(m)$ is possible.

Theorem 6 is proved by precisely similar arguments but using Theorem 4 instead of Theorem 3.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

[8] E. T. Bell, *loc. cit.*

NEW RESULTS FOR THE NUMBER $g(n)$ IN WARING'S PROBLEM.[1]

By H. S. Zuckerman.

I. *Introduction.* Waring's problem deals with the representation of positive integers by sums of positive or zero integral n-th powers. The number $g(n)$ is defined as the minimum s such that

$$(1) \qquad m = \sum_{i=1}^{s} x_i^n, \quad x_i \text{ integers}, \quad x_i \geqq 0, \qquad (i = 1, 2, \cdots, s)$$

has at least one solution $x_1, x_2, \cdots, x_s$ for every integer $m > 0$. The number $G(n)$ is the minimum s such that (1) has at least one solution for every sufficiently large integer m. We shall define $G(n, c)$ as the minimum s such that (1) has at least one solution for every integer $m \geqq c$. Then

$$g(n) = G(n, 1), \quad G(n, c+1) \leqq G(n, c), \quad G(n) = \lim_{c=\infty} G(n, c).$$

A decomposition of m is a representation of m as a sum of positive integral n-th powers, in the form (1) with the added restriction $x_i > 0$. The number of n-th powers, s, is the weight of the decomposition.

Let us consider the m in the range $0 < m < 3^n$. Write

$$3^n = q2^n + r, \quad q, r \text{ integers}, \quad 0 < r < 2^n.$$

In any decomposition of m each x_i is either 2 or 1. Then the decompositions of m of minimum weight are given by

$$m = y2^n + z, \quad 0 \leqq z < 2^n \text{ for } 0 \leqq y < q, \quad 0 \leqq z < r \text{ for } y = q$$

of weight $y + z$. Since $r \leqq 2^n - 1$ the maximum weight for any m in the range $0 < m < 3^n$ is

$$I = q + 2^n - 2,$$

which is the weight of the decomposition of $q2^n - 1$. Therefore

$$g(n) \geqq I = q + 2^n - 2$$

and, since 3^n is a decomposition of 3^n of weight 1, I n-th powers suffice to represent all integers in the range $0 < m \leqq 3^n$.

[1] Presented to the American Mathematical Society April 6, 1935 and December 31, 1935.

I. Vinogradow[2] has proved that

$$G(n) \leqq [n(6 \log n + \log 216 + 4)], \; n > 3$$

by proving the existence of a number V_n such that

$$(2) \qquad G(n, V_n) \leqq [n(6 \log n + \log 216 + 4)].$$

In parts IV and V of the author's doctor's dissertation at the University of California the Vinogradow proof is reproduced and the constants determined as they occur. In some cases proofs are supplied for results which Vinogradow states without proof. In part V we conclude that (2) holds if

$$(3) \qquad \log V_n = 18n2^{14n^2/3}, \; n > 4.$$

The upper bound (2) for $G(n, V_n)$ is less than I for $n > 7$.

In part II of this paper we develop a method for dealing with the integers between 3^n and V_n without the use of tables. In part III we apply this method for some values of n to obtain upper bounds for $g(n)$ using the results of part V of the author's dissertation and the fact that I n-th powers suffice to represent all integers between zero and 3^n. In the cases where this upper bound is equal to I we then have $g(n) = I$. It has been known for a long time that $g(n) = I$ when $n = 2$ or 3 but this is the first time the formula $g(n) = I$ has been proved for other values of n. The results are tabulated at the end of part III.

II. *Decomposition of integers between 3^n and V_n.* We define the integers q, r, u, v, w, by

$$(4) \qquad \begin{gathered} 3^n = q2^n + r, \quad 4^n = u3^n + v2^n + w, \\ 0 < r < 2^n, \quad 0 < w < 2^n, \quad 0 \leqq v \leqq q, \quad u = [4^n/3^n]. \end{gathered}$$

Then we have

$$(5) \qquad \begin{aligned} (k - ax)4^n + (l + uax + bx)3^n + (m + vax - qbx)2^n \\ = k4^n + l3^n + m2^n + (br - aw)x. \end{aligned}$$

We take k, l, m, a, b, x to be integers, and $x \geqq 0$. If $k \geqq ax$, $l \geqq - uax - bx$, $m \geqq - vax + qbx$ the left member of (5) is a decomposition of $k4^n + l3^n + m2^n + (br - aw)x$ of weight

[2] I. Vinogradow, "On Waring's problem," *Annals of Mathematics*, vol. 36 (1935), pp. 395-405.

$$k+l+m-ax+uax+bx+vax-qbx.$$

If $br>aw$ we can obtain a decomposition of any integer between

$$k4^n+l3^n+m2^n+(br-aw)x \quad \text{and} \quad k4^n+l3^n+m2^n+(br-aw)(x+1)-1$$

inclusive by adding 1^n at most $br-aw-1$ times to the decomposition of $k4^n+l3^n+m2^n+(br-aw)x$. Then the highest weight in this interval is

$$k+l+m+(-a+ua+b+va-qb)x+br-aw-1.$$

We now let x take the values $0, 1, \cdots, [2^n/(br-aw)]$ and obtain decompositions for every integer between

$$k4^n+l3^n+m2^n \quad \text{and} \quad k4^n+l3^n+(m+1)2^n+\phi$$

inclusive, where

$$\phi=[2^n/(br-aw)](br-aw)+br-aw-1-2^n \geqq 0.$$

The greatest weight in this interval is

$$k+l+m+\max(0,-a+ua+b+va-qb)[2^n/(br-aw)]+br-aw-1.$$

If $br<aw$ we can obtain a decomposition of any integer between

$$k4^n+l3^n+m2^n-(aw-br)x \quad \text{and} \quad k4^n+l3^n+m2^n-(aw-br)(x-1)-1$$

inclusive by adding 1^n at most $aw-br-1$ times to our decomposition of

$$k4^n+l3^n+m2^n+(br-aw)x.$$

Then the highest weight in this interval is

$$k+l+m+(-a+ua+b+va-qb)x+aw-br-1.$$

We now let x take the values $0, 1, \cdots, [2^n/(aw-br)]$ and obtain decompositions for every integer between

$$k4^n+l3^n+(m-1)2^n+\psi \quad \text{and} \quad k4^n+l3^n+m2^n+aw-br-1$$

inclusive, where

$$\psi=2^n-[2^n/(aw-br)](aw-br) \leqq aw-br-1.$$

The greatest weight in this interval is

$$k+l+m+\max(0,-a+ua+b+va-qb)[2^n/(aw-br)]+aw-br-1.$$

In both cases we must have $k \geqq ax$, $l \geqq -uax - bx$, $m \geqq -vax + qbx$. Hence we take

$$k = \max(0, a)[2^n/|br - aw|], \quad l = \max(0, -ua - b)[2^n/|br - aw|],$$
$$m = \max(0, qb - va)[2^n/|br - aw|].$$

Let W be the greatest weight in the interval. Then, in both cases, we have

$$W = \{\max(0, a) + \max(0, -ua - b) + \max(0, qb - va)$$
$$+ \max(0, -a + ua + b + va - qb)\}[2^n/|br - aw|] + |br - aw| - 1.$$

We proceed to show that changing the signs of both a and b leaves W and the interval invariant. Let W_1 be the value of W for $a = A$, $b = B$ and let W_2 be its value for $a = -A$, $b = -B$. By choice of notation we can make $Br > Aw$. Let L_1 be the lower end of the interval for $a = A$, $b = B$; L_2 for $a = -A$, $b = -B$. Let U_1 and U_2 be the two upper ends of the intervals. Then

$$W_1 = \{\max(0, A) + \max(0, -uA - B) + \max(0, qB - vA)$$
$$+ \max(0, -A + uA + B + vA - qB)\}[2^n/(Br - Aw)] + Br - Aw - 1,$$
$$W_2 = \{\max(0, -A) + \max(0, uA + B) + \max(0, -qB + vA)$$
$$+ \max(0, A - uA - B - vA + qB)\}[2^n/(Br - Aw)] + Br - Aw - 1.$$

Since $\max(0, \alpha) - \max(0, -\alpha) = \alpha$, we have

$$W_1 - W_2 = (A - uA - B + qB - vA - A + uA + B + vA - qB)$$
$$\times [2^n/(Br - Aw)] = 0.$$

Also, we have

$$L_1 = \max(0, A)[2^n/(Br - Aw)]4^n + \max(0, -uA - B)[2^n/(Br - Aw)]3^n$$
$$+ \max(0, qB - vA)[2^n/(Br - Aw)]2^n,$$
$$L_2 = \max(0, -A)[2^n/(Br - Aw)]4^n + \max(0, uA + B)[2^n/(Br - Aw)]3^n$$
$$+ \max(0, -qB + vA)[2^n/(Br - Aw)]2^n - 2^n + \psi,$$
$$\psi = 2^n - [2^n/(Br - Aw)](Br - Aw),$$

and hence

$$L_1 - L_2 = \{A4^n + (-uA - B)3^n + (qB - vA)2^n\}[2^n/(Br - Aw)]$$
$$+ 2^n - 2^n + [2^n/(Br - Aw)](Br - Aw) = 0$$

by (4). Finally, we have

$$U_1 = L_1 + 2^n + \phi,$$
$$U_2 = L_2 + 2^n - Aw + Br - 1 - \psi,$$
$$\phi = [2^n/(Br - Aw)](Br - Aw) + Br - Aw - 1 - 2^n,$$
$$U_1 - U_2 = L_1 - L_2 + \phi + Aw - Br + 1 + \psi$$
$$= [2^n/(Br - Aw)](Br - Aw) + Br - Aw - 1 - 2^n + Aw - Br + 1 +$$
$$- [2^n/(Br - Aw)](Br - Aw) = 0.$$

Therefore $L_1 = L_2$, $U_1 = U_2$, $W_1 = W_2$ so that the value of W and the interval remain unchanged when we change the signs of a and b. Hence we can, without loss of generality, choose a and b so that $br > aw$. Then

$$W = \{\max(0, a) + \max(0, -ua - b) + \max(0, qb - va) + \max(0, -a + ua + b + va - qb)\}[2^n/(br - aw)] + br - aw - 1, \tag{6}$$

$$k = \max(0, a)[2^n/(br - aw)], \quad l = \max(0, -ua - b)\ [2^n/(br - aw)] \quad m = \max(0, qb - va)[2^n/(br - aw)], \tag{7}$$

and the interval is

$$k4^n + l3^n + m2^n \text{ to } k4^n + l3^n + (m+1)2^n + \phi \text{ inclusive}, \tag{8}$$

$$\phi = [2^n/(br - aw)](br - aw) + br - aw - 1 - 2^n. \tag{9}$$

For each value of n we choose values for a and b to make W small and then determine the interval over which W suffices by (7), (8) and (9). Call this interval I_2. We then make ascents from I_2 using the two following theorems proved by L. E. Dickson.[3]

THEOREM 1. *If every integer $> l$ and $\leqq g$ is a sum of $k - 1$ integral n-th powers $\geqq 0$, and if m is an integer for which*

$$(m+1)^n - m^n < g - l$$

then every integer $> l$ and $\leqq g + (m+1)^n$ is a sum of k integral n-th powers $\geqq 0$.

THEOREM 2. *Let l be an integer $\geqq 0$. Let*

$$\nu = (1 - l/L_0)/n, \quad L_0 > l, \quad (\nu L_0)^{n/(n-1)} \geqq L_0.$$

Compute L_t by $\log L_t = (n/(n-1))^t(\log L_0 + n \log \nu) - n \log \nu$. If all integers between l and L_0 inclusive are sums of k integral n-th powers $\geqq 0$, then all integers between l and L_t inclusive are sums of $k + t$ integral n-th powers $\geqq 0$.

To take care of the integers between 3^n and I_2 we use either a result proved by L. E. Dickson[4] or, if possible, choose a pair of values for a and b so that the interval I_1 includes the number 3^n and then ascend from I_1 to I_2.

---

[3] L. E. Dickson, "Recent progress on Waring's theorem and its generalizations," *Bulletin of the American Mathematical Society*, vol. 39 (1933), pp. 701-727, Theorems 10 and 12.

[4] *Ibid.*, Theorems 5 and 6.

III. *The special case* $n = 15$. We shall now apply the method of part II to the case $n = 15$. We have

$$2^n = 32\ 768, \quad 3^n = 14\ 348\ 907, \quad 4^n = 1\ 073\ 741\ 824.$$

By division we obtain

$$q = 437, \quad r = 29291, \quad u = 74, \quad v = 363, \quad w = 27922,$$
$$I = q + 2^n - 2 = 33203.$$

Taking $a = 0$, $b = 1$ we find, by (6), (7), (8), and (9),

$$W = 29727, \quad I_1: \text{ from } 3^n - 29291 \text{ to } 3^n + 29290.$$

Taking $a = 1$, $b = 1$ we find

$$W = 3093, \quad I_2: \text{ from } 23 \cdot 4^n + 3 \cdot 3^n + 388 \cdot 2^n + 10431 \text{ to }$$
$$23 \cdot 4^n + 3 \cdot 3^n + 389 \cdot 2^n + 10518.$$

That is, 29727 fifteenth powers suffice to represent every integer between $3^n - 29291$ and $3^n + 29290$ and 3093 suffice to represent every integer between

$$23 \cdot 4^n + 3 \cdot 3^n + 388 \cdot 2^n + 10431 \quad \text{and} \quad 23 \cdot 4^n + 3 \cdot 3^n + 389 \cdot 2^n + 10518.$$

Starting from I_1 we ascend by Theorem 1, applying it 437 times with $m = 1$, 74 times with $m = 2$, and 23 times with $m = 3$. Then

$$29727 + 437 + 74 + 23 = 30261$$

fifteenth powers suffice to represent every integer from I_1 to I_2.

Ascending from I_2 we apply Theorem 1, 436 times with $m = 1$; 73 times with $m = 2$; 27 times with $m = 3$; 14 times with $m = 4$; 9 times with $m = 5$; 6 times with $m = 6$; 5 times with $m = 7$; 4 times with $m = 8$; 3 times with $m = 9$; and 2 times each with $m = 10, 11, 12, 13, 14$. We have then applied the theorem 587 times. Hence $3093 + 587 = 3680$ fifteenth powers suffice to represent every integer from

$$23 \cdot 4^n + 3 \cdot 3^n + 388 \cdot 2^n + 10431 \text{ to }$$
$$2 \cdot 15^n + 23 \cdot 4^n + 3 \cdot 3^n + 388 \cdot 2^n + 10431.$$

We now apply Theorem 2 with

$$l = 23 \cdot 4^n + 3 \cdot 3^n + 388 \cdot 2^n + 10431, \quad L_0 = l + 2 \cdot 15^n, \quad k = 3680,$$

and

$$\nu = (1 - (l/L_0))/15 = 0.066.$$

Then

$$\log L_t = (15/14)^t(\log L_0 + 15 \cdot \log \nu) - 15 \cdot \log \nu$$
$$\log \log L_t > .02996t - .521 \text{ (logarithms to the base 10)}$$

and all integers between l and L_t are sums of $3680 + t$ integral fifteenth powers $\geqq 0$. If we choose $t = I - k = 29523$ we have

$$\log \log L_t > 883 \text{ (logarithms to the base 10)}.$$

In the introduction we showed that I n-th powers suffice to represent every positive integer $\leqq 3^n$. Hence, since I_1 includes 3^n, we have that every integer m,

$$0 < m \leqq 10^{10^{883}},$$

is a sum of $I = 33203$ integral fifteenth powers $\geqq 0$.

By the results of part V of the author's dissertation we have the following inequality for the Vinogradow constant when $n = 15$.

$$\log \log V_{15} < 319 \text{ (logarithms to the base 10)}.$$

If we choose $t = 10665$ we have

$$\log \log L_t > 319 \text{ (logarithms to the base 10)}$$

so all integers between l and V_{15} are sums of $10665 + 3680 = 14345$ integral fifteenth powers $\geqq 0$. For $n = 15$ the inequality (2) is

$$G(15, V_{15}) \leqq 384.$$

We now have the following results.

$$G(15, 10^{10^{319}}) \leqq 384, \qquad G(15, 24 \cdot 4^n) \leqq 14345,$$
$$G(15, 3^n) \leqq 30261, \qquad g(15) \leqq 33203.$$

The second inequality follows from the facts that all integers between l and V_{15} are sums of 14345 and all integers greater than V_{15} are sums of 384. We have replaced the value of l by a larger value.

Since $I = 33203$ we have $g(15) = I = 33203$.

Let L be the value of L_t obtained by setting $t = I - k$. By treating other values of n as we did $n = 15$ we obtain the following results. In the cases $n = 17$ and $n = 20$ we cannot find a suitable interval I_1 so we use L. E. Dickson's result referred to above.

List of results for various n from 14 to 20.

$$n = 14.$$

Use $a = 0,\ b = 1$ and $a = -1,\ b = -1$.

$$\log\log L > 249,\quad \log\log V_{14} < 278,\quad g(14) \leqq 17555,\quad g(14) \geqq I = 16673,$$
$$G(14, V_{14}) \leqq 352.$$

$$n = 15.$$

Use $a = 0,\ b = 1$ and $a = 1,\ b = 1$.

$$\log\log L > 883,\quad \log\log V_{15} < 319,\quad g(15) = I = 33203,$$
$$G(15, 3^n) \leqq 30261,\quad G(15, 24 \cdot 4^n) \leqq 14345,\quad G(15, V_{15}) \leqq 384.$$

$$n = 16.$$

Use $a = 0,\ b = 1$ and $a = 1,\ b = 1$.

$$\log\log L > 1630,\quad \log\log V_{16} < 363,\quad g(16) = I = 66190,$$
$$G(16, 3^n) \leqq 56526,\quad G(16, 12 \cdot 4^n) \leqq 20974,\quad G(16, V_{16}) \leqq 416.$$

$$n = 17.$$

Use $a = 0,\ b = 1$.

$$\log\log L > 2464,\quad \log\log V_{17} < 409,\quad g(17) = I = 132055,$$
$$G(17, 3 \cdot 3^n) \leqq 53984,\quad G(17, V_{17}) \leqq 448.$$

$$n = 18.$$

Use $a = 0,\ b = 1$ and $a = 2,\ b = 1$.

$$\log\log L > 2472,\quad \log\log V_{18} < 459,\quad g(18) = I = 263619,$$
$$G(18, 3^n) \leqq 236932,\quad G(18, 3 \cdot 4^n) \leqq 182522,\quad G(18, V_{18}) \leqq 480.$$

$$n = 19.$$

Use $a = 0,\ b = 1$ and $a = 1,\ b = 1$.

$$\log\log L > 5339,\quad \log\log V_{19} < 511,\quad g(19) = I = 526502,$$
$$G(19, 3^n) \leqq 443926,\quad G(19, 2 \cdot 4^n) \leqq 320864,\quad G(19, V_{19}) \leqq 513.$$

$$n = 20.$$

Use $a = 1,\ b = 1$.

$$\log\log L > 20042,\quad \log\log V_{20} < 566,\quad g(20) = I = 1051899,$$
$$G(20, 8 \cdot 4^n) \leqq 177571,\quad G(20, V_{20}) \leqq 546.$$

The logarithms above are all to base 10.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

ON WARING'S PROBLEM WITH POLYNOMIAL SUMMANDS.

By LOO-KENG HUA.[1]

1. *Introduction and notation.* Let

$$P(h) = \alpha_0 h^k + \alpha_1 h^{k-1} + \cdots + \alpha_k$$

be an integral-valued polynomial with $\alpha_0 > 0$;

$r_{P,s}(n) = r(n) =$ the number of solutions of the diophantine equation

(1) $\quad n = P(h_1) + P(h_2) + \cdots + P(h_s), \quad h_i \geqq 0 \quad (i = 1, 2, \cdots, s);$

$k =$ integer $\geqq 3$; $\; s \geqq K(k-2) + 5$; $\; K = 2^{k-1}$;

$d =$ the least common denominator of $\alpha_0, \alpha_1, \cdots, \alpha_k$;
the canonical product of $n = p_1^{l_1} \cdots p_t^{l_t}$;
$\theta_i =$ a positive integer or zero such that $p_i^{\theta_i} \mid d$, $p_i^{\theta_i+1} \nmid d$;

$$n^* = p_1^{l_1+\theta_1} \cdots p_t^{l_t+\theta_t};$$
$$\rho = e^{2\pi i (l/q)}, \qquad (l, q) = 1;$$
$$S_\rho = \sum_{z=1}^{q^*} \rho^{P(z)};$$

$A_{P,s}(q, j) = A(q) = (1/q^{*s}) \sum_\rho S_\rho^s \rho^{-j}$, where the summation is taken over all q-th primitive roots ρ of unity;

$$\mathfrak{S}(j, P, s, \lambda) = \mathfrak{S}(j, \lambda) = \sum_{q=1}^{\lambda} A_{P,s}(q, j);$$
$$\mathfrak{S}_j = \mathfrak{S}(j, \infty);$$

$j_w =$ the greatest real root of $P(x) = j$ if there is any;

$$f(x) = \sum_{h=0}^{n_w} x^{P(h)};$$
$$\psi_\rho(x) = \frac{\Gamma(1+a)}{\alpha_0^a} \frac{S_\rho}{q^*} \left(1 + \sum_{j=1}^{n} \frac{a(a+1)\cdots(a+j-1)}{j!} (x/\rho)^j\right);$$
$$\phi_\rho(x) = \frac{\Gamma(1+a)}{\alpha_0^a} \frac{S_\rho}{q^*} \left(\sum_{j=n+1}^{\infty} \frac{a(a+1)\cdots(a+j-1)}{j!} (x/\rho)^j\right);$$
$$\Psi_\rho(x) = \psi_\rho(x) + \phi_\rho(x) = \frac{P(1+a)}{\alpha_0^a} \frac{S_\rho}{q^*} (1 - x/\rho)^a.$$

---

[1] Research fellow of the China Foundation for the Promotion of Education and Culture.

Throughout the paper ϵ denotes an arbitrarily small positive number and $C_1, C_2, \cdots$, denote positive numbers depending at most on P, s and ϵ.

It is to be noticed that if $P(x)$ is a polynomial with integral coefficients, then $\mathfrak{S}_j$ reduces to that treated by Landau[2] (1), (2).

We shall prove the following result:

THEOREM 1. *We have*

$$| r_{P,s}(n) - \alpha_0^{-sa} \frac{P^s(1+a)}{P(sa)} \mathfrak{S}_n n^{sa-1} | < C_1 n^{sa-1-a/K+\epsilon}.$$

The method used is due to Gelbcke (3) for $P(x) = x^k$. We shall omit proofs of results when the method of proof differs only slightly from that given by Gelbcke.

2. *Lemmas for polynomial $P(x)$.*

LEMMA 1. *The necessary and sufficient condition that a polynomial $P(x)$ of k-th degree in the rational field be integral-valued is*

$$P(x) = a_0 P_k(x) + \cdots + a_k P_0(x) \tag{2}$$

where

$$P_i(x) = \frac{x(x-1)\cdots(x-i+1)}{i!}, \qquad (i = 1, \cdots, k);\ P_0(x) = 1,$$

and the a_i's are integers. [Hilbert (7)].

LEMMA 2. *The necessary and sufficient condition that there do not exist integers u and l (> 1) such that*

$$P(x) \equiv u \pmod{l}$$

for all values of x, is

$$(a_0, a_1, \cdots, a_{k-1}) = 1,$$

where $P(x)$ is given by (2).

LEMMA 3. *If all sufficiently large integers are sums of s values of $P(h)$, $h \geqq 0$, then they are also sums of s values of the polynomial $Q(h) = P(h-t)$, where t is a positive integer.*

Proof. If

$$n = P(h_1) + \cdots + P(h_s), \qquad h_i \geqq 0,$$

then evidently

$$n = Q(h_1 + t) + \cdots + Q(h_s + t), \qquad h_i + t \geqq 0.$$

[2] See the list of references at the end of the paper, page 562.

LEMMA 4. *There exists a positive integer t such that the coefficients of $P(x+t)$ are all positive.*

It follows from Lemmas 3 and 4 that the study of Waring's problem for a general integral-valued polynomial $P(h)$ depends on the same problem for a polynomial with positive coefficients.

Henceforth we shall assume that all α's are positive and the constant term is zero. Under such an assumption j_w is uniquely determined and is greater than zero for $j > 0$.

3. *Well known lemmas.*

LEMMA 5. *Let $|y| \leqq 1/2$ and $a_0 \geqq a_1 \geqq \cdots \geqq 0$, then*

$$\left| \sum_{j=1}^{N} a_j e^{2\pi i y j} \right| \leqq \frac{a_0}{\sin \pi |y|}.$$

[Landau (6), Theorem 140 with $R(w) = \sum_{j=1}^{w} e^{2\pi i y j}$].

LEMMA 6. *We have*

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \sum_{j=0}^{N} a_j e^{2\pi i y j} \right|^2 dy = \sum_{j=0}^{N} |a_j|^2.$$

[Landau (6), Theorem 223].

LEMMA 7. *For $\beta > 0$ and j an integer > 0,*

$$\left| \frac{\Gamma(1+\beta+j)}{j!} - j^{\beta} \right| < \gamma(\beta) j^{\beta-1},$$

where $\gamma(\beta)$ depends on β only.

LEMMA 8. *We have*

$$\sum_{j=0}^{w} r_2{}^2(j) < C_2 w^{2a+\epsilon}.$$

Proof. Landau, *Mathematische Zeitschrift*, vol. 31 (1929-30), p. 149.

LEMMA 9. *Let ρ be a primitive q-th root of unity, $m > 0$, r integers, then*

$$\left| \sum_{h=r+1}^{r+m} \rho^{P(h)} \right|^K < C_6 q^{\epsilon} m^{\epsilon} (m^{K-1} + m^K/q + q m^{K-k}).$$

Proof. By the same method used by Landau (6), Theorem 267, we can prove that

$$\left| \sum_{h=r+1}^{r+m} \rho^{Q(h)} \right|^K < C_7 q^{\epsilon} m^{\epsilon} (m^{K-1} + m^K/q + q m^{K-k}),$$

where $Q(h)$ is a polynomial with real coefficients and its first coefficient $= 1$.

Suppose that

$$\rho^{a_0} = e^{2\pi i(l/q)a_0} = e^{2\pi i(la_0/qk!)}$$

is a primitive q'-th root of unity. If $(l, k!) = 1$, then also $(l, qk!) = 1$ and

$$\begin{aligned} q' = qk! \qquad &\text{if} \quad (a_0, qk!) = 1, \\ \left.\begin{aligned} q' &< qk! \\ q' &> (qk!)/a_0 \end{aligned}\right\} \qquad &\text{if} \quad (a_0, qk!) > 1. \end{aligned}$$

From the first part we have

$$\begin{aligned} \Big| \sum_{h=r+1}^{r+m} \rho^{P(h)} \Big|^K &< C_7 q'^{\epsilon} m^{\epsilon} (m^{K-1} + m^K/q' + q' m^{K-k}) \\ &< C_6 q^{\epsilon} m^{\epsilon} (m^{K-1} + m^K/q + q m^{K-k}). \end{aligned}$$

If $(l, k!) \neq 1$, then there exists an integer t such that $(l + tq, k!) = 1$. In fact, since $(l, q) = 1$, the arithmetic progression $l + tq$ contains a prime which is not a divisor of k.

Therefore

$$\begin{aligned} \Big| \sum_{h=r+1}^{r+m} e^{2\pi i(l/q)P(h)} \Big|^K &= \Big| \sum_{h=r+1}^{r+m} e^{2\pi i[(l+qt)/q]P(h)} \Big|^K \\ &< C_6 q^{\epsilon} m^{\epsilon} (m^{K-1} + m^K/q + q m^{K-k}), \end{aligned}$$

since $P(x)$ is an integral-valued polynomial and therefore $e^{2\pi i t P(h)} = 1$.

LEMMA 10. *We have*

$$| S_\rho | < C_8 q^{1-1/K+\epsilon}.$$

LEMMA 11. *We have*

$$| A(q) | < C_1 q^{-(5/4)+\epsilon}.$$

Proof. These two lemmas follow at once from Lemma 9 in the same way that Landau's Theorems 269 and 270 follow from his Theorem 267. We use the fact that $q \leqq q^* \leqq k!q$.

5. *Farey cut.* On the unit circle $|x| = 1$, we take points $\rho = e^{2\pi i(l/q)}$ which correspond to Farey fractions l/q with denominators less than or equal to $n^{1-\alpha}$, and divide the circumference into sub-arcs by means of mediants. Let us write

$$x = \rho e^{2\pi i y}.$$

Then on the arc associated with ρ, we shall have

$$-y_1 \leqq y \leqq y_2,$$

where

$$\frac{1}{2qn^{1-a}} \leqq y_1 < \frac{1}{qn^{1-a}}, \qquad \frac{1}{2qn^{1-a}} \leqq y_2 < \frac{1}{qn^{1-a}}.$$

We consider minor and major arcs given by

minor arcs $\mathfrak{m}$: $\quad n^a \leqq q \leqq n^{1-a},$

major arcs $\mathfrak{M}$: $\quad 1 \leqq q < n^a.$

Each arc $\mathfrak{M}$ is further divided into two parts. That part of $\mathfrak{M}$ for which

$$|y| \leqq \frac{1}{2q^{1/2}n^{1-a/2}}$$

is called an arc of type $\mathfrak{M}_1$ and the remaining part of $\mathfrak{M}$ in which

$$|y| > \frac{1}{2q^{1/2}n^{1-a/2}}$$

is called an arc of type $\mathfrak{M}_2$. There are always arcs of type $\mathfrak{M}_2$, since

$$\frac{1}{2q^{1/2}n^{1-a/2}} \leqq \frac{1}{2qn^{1-a}},$$

where $q < n^a$.

6. *Further lemmas.*

LEMMA 12. *On the whole circle $|x| = 1$ we have*

$$|\psi_\rho(x)| < C_{10}n^a q^{-1/K+\epsilon}.$$

Proof. From the definition of $\psi_\rho(x)$ we have

$$\psi_\rho(x) = \frac{S_\rho}{\alpha_0{}^a q^*}\left(\Gamma(1+a) + a\sum_{j=1}^{n}\frac{\Gamma(1+a+j)}{j!\,\Gamma(a+j)}(x/\rho)^j\right).$$

From Lemmas 10 and 7, it follows that

$$\begin{aligned}|\psi_\rho(x)| &< C_8 q^{-1/K+\epsilon}(\Gamma(1+a) + a\sum_{j=1}^{n}(j^{a-1}+\gamma(a)j^{a-2})) \\ &< C_{10}n^a q^{-1/K+\epsilon}.\end{aligned}$$

LEMMA 13. *If* $|y| \leqq 1/2$,

then

$$|\phi_\rho(x)| < C_{11}n^{a-1}q^{-1/K+\epsilon}|y|^{-1},$$
$$|\Psi_\rho(x)| < C_{12}q^{-1/K+\epsilon}|y|^{-a},$$

and

$$|\psi_\rho(x)| < C_{13}q^{-1/K+\epsilon}\,\mathrm{Min}(n^a, |y|^{-a}).$$

The proof of these inequalities is the same as that of Theorems 7, 8 of Gelbcke (3).

LEMMA 14. *We have*

$$\sum_{\mathfrak{M}_1} \int_{\mathfrak{K}-\mathfrak{M}_1} | \psi_\rho{}^s(x) | \, | dx | < C_{14} n^{sa-1-(s/K-2)a+\epsilon},$$

where $\psi_\rho(x)$ *corresponds to the arc* $\mathfrak{M}_1$, *and the path of integration is the whole circle excluding the arc* $\mathfrak{M}_1$.

The proof of this lemma is similar to that of Theorem 9 of Gelbcke (3).

7. *Lemmas for the arcs* $\mathfrak{m}$.

LEMMA 15. *On* $\mathfrak{m}$ *we have*

$$| f(x) | < C_{15} n^{a-a/K+\epsilon}.$$

Proof. Let

$$\tau(j) = \sum_{h=0}^{j_w} \rho^{P(h)} \quad \text{for} \quad j \geqq 0, \qquad \tau(-1) = 0.$$

Then

$$\begin{aligned} f(x) &= \sum_{h=0}^{n_w} \rho^{P(h)} (x/\rho)^{P(h)} \\ &= \sum_{j=0}^{n} [\tau(j) - \tau(j-1)] (x/\rho)^j \\ &= \sum_{j=0}^{n-1} \tau(j) [(x/\rho)^j - (x/\rho)^{j+1}] + \tau(n)(x/\rho)^n \\ &= (1 - x/\rho) \sum_{j=0}^{n-1} \tau(j)(x/\rho)^j + \tau(n)(x/\rho)^n. \end{aligned}$$

Now

$$| \arg (x/\rho) | < 2\pi/qn^{1-a},$$

therefore the length of the chord

$$| 1 - x/\rho | < 2\pi/qn^{1-a}.$$

Further from Lemma 9 it follows that

$$\begin{aligned} | \tau(j) |^K &< C_6 q^\epsilon n_w{}^\epsilon (n_w{}^{K-1} + n_w{}^K/q + qn_w{}^{K-k}) \\ &< C_{16} q^\epsilon n^{a\epsilon} (n^{a(K-1)} + n^{aK}/q + qn^{a(K-k)}) \end{aligned}$$

since $j_w < C_{17} j^a$. Then

$$| \tau(j) | < C_{18} n^{a-a/K+\epsilon}.$$

Therefore

$$| f(x) | < [(2\pi/qn^{1-a})n + 1] \, C_{18} n^{a-a/K+\epsilon} < C_{15} n^{a-aK/+\epsilon}.$$

LEMMA 16. *We have*

$$\sum_{\mathfrak{m}} \int_{\mathfrak{m}} | f^s(x) | \; | dx | < C_{19} n^{(s-4)(a-a/K)+2a+\epsilon}.$$

Proof. See Gelbcke (3).

8. *Lemmas for the arcs* $\mathfrak{M}_1$.

LEMMA 17. *If* $| y | \leqq 1/2$, *then*

$$| f(x) - \psi_\rho(x) | < C_{22} q^{1-1/K+\epsilon} \operatorname{Max}(n | y |, 1).$$

Proof. As in Lemma 15 we may write

$$\begin{aligned} f(x) &= (1 - x/\rho) \sum_{j=1}^{n-1} \tau(j)(x/\rho)^j + \tau(n)(x/\rho)^n, \\ \psi_\rho(x) &= \frac{\Gamma(1+a)}{\alpha_0{}^a} \frac{S_\rho}{q^*} \left(1 + \sum_{j=1}^{n} \frac{a(a+1)\cdots(a+j-1)}{j!} (x/\rho)^j \right) \\ &= \frac{S_\rho}{\alpha_0{}^a q^*} \left((1 - x/\rho) \sum_{j=1}^{n-1} \frac{\Gamma(1+a+j)}{j!} (x/\rho)^j + \frac{\Gamma(1+a+n)}{n!} (x/\rho)^n \right) \\ &= (1 - x/\rho) \sum_{j=1}^{n-1} v(j)(x/\rho)^j + v(n)(x/\rho)^n, \end{aligned}$$

where

$$v(j) = \frac{S_\rho}{\alpha_0{}^a q^*} \frac{\Gamma(1+a+j)}{j!}, \qquad (j = 1, \cdots, n).$$

For every $j \geqq 0$, $\tau(j)$ has $[j_w] + 1$ terms. Divide $\tau(j)$ into partial sums, each $= S_\rho$, plus $[j_w] + 1 - [j_w/q^*]q^* \leqq q^*$ terms. By Lemma 9, we have

$$| \tau(j) - [j_w/q^*] S_\rho | < C_{23} q^{1-1/K+\epsilon},$$
$$| \tau(j) - (j_w/q^*) S_\rho | < C_{23} q^{1-1/K+\epsilon} + | S | < C_{24} q^{1-1/K+\epsilon}.$$

Now

$$\left| \frac{\Gamma(1+a+j)}{j!} - j^a \right| < C_{25} j^{a-1} < C_{26},$$

and

$$\begin{aligned} | j^a - \alpha_0{}^a j_w | &= | (\alpha_0 j_w{}^k + \cdots + \alpha_k)^a - \alpha_0{}^a j_w | \\ &\leqq j_w | (\alpha_0 + \alpha_1/j_w + \cdots + \alpha_k/j_w{}^k)^a - \alpha_0{}^a | < C_{27}. \end{aligned}$$

Therefore

$$| v(j) - j_w(S_\rho/q^*) | < C_{28} | S_\rho/q^* | < C_{29}.$$

Thus we have

$$| \tau(j) - v(j) | < C_{30} q^{1-1/K+\epsilon},$$

and consequently

$$|f(x)-\psi_\rho(x)| = |(1-x/\rho)\sum_{j=0}^{n-1}[\tau(j)-v(j)](x/\rho)^j + [\tau(n)-v(n)](x/\rho)^n|$$
$$< C_{31}(n|y|+1)q^{1-1/K+\epsilon} < C_{22}q^{1-1/K+\epsilon}\operatorname{Max}(n|y|,1).$$

LEMMA 18. *We have*

$$\sum_{\mathfrak{M}_1}\int_{\mathfrak{M}_1}|f^s(x)-\psi_\rho{}^s(x)|\,|dx| < C_{32}n^{sa-1-(s/K-2)a+\epsilon}$$
$$+C_{33}n^{sa-1-a+\epsilon}\begin{cases} n^{(3/4)a} & \text{for } k=3,\\ n^{(3/8)a} & \text{for } k=4,\\ 1 & \text{for } k=5.\end{cases}$$

Proof. See Gelbcke (3).

9. *Lemmas for the arcs* $\mathfrak{M}_2$.

LEMMA 19. *If* $|y|\leqq 1/2$ *then for* ρ *corresponding to* $\mathfrak{M}_2$ *we have*

$$|f(x)| < C_{39}n^{a-1/K+\epsilon}q^{-1/K-\epsilon}|y|^{-1/K-\epsilon}.$$

Proof. See Gelbcke (3).

LEMMA 20. *We have*

$$\sum_{\mathfrak{M}_2}\int_{\mathfrak{M}_2}|f^s(x)|\,|dx| < C_{43}n^{sa-1-(s/K-1)(a/2)+\epsilon}\begin{cases} n^{(3/8)a} & \text{for } k=3,\\ n^{(3/4)a} & \text{for } k=4,\\ 1 & \text{for } k\geqq 5.\end{cases}$$

Proof. See Gelbcke (3).

10. *Proof of Theorem* 1. It is evident that the coefficient of the n-th power of x in $f^s(x)$ is equal to $r_{P,s}(n)$. By Cauchy's Theorem we have

$$r_{P,s}(n) = \frac{1}{2\pi i}\int_{\mathfrak{K}}\frac{f^s(x)}{x^{n+1}}\,dx$$
$$=\frac{1}{2\pi i}\left(\sum_{\mathfrak{M}}\int_{\mathfrak{M}}\frac{f^s(x)}{x^{n+1}}\,dx + \sum_{\mathfrak{M}_2}\int_{\mathfrak{M}_2}\frac{f^s(x)}{x^{n+1}}\,dx + \sum_{\mathfrak{M}_1}\int_{\mathfrak{M}_1}\frac{f^s(x)}{x^{n+1}}\,dx\right).$$

By Lemmas 16, 20, we have

$$(5)\qquad \left|r_{P,s}(n)-\sum_{\mathfrak{M}_1}\int_{\mathfrak{M}_1}\frac{f^s(x)}{x^{n+1}}\,dx\right| < C_{19}n^{(s-4)(a-a/K)+2a+\epsilon}$$
$$+C_{43}n^{sa-1-(s/K-1)(a/2)+\epsilon}\begin{cases} n^{(3/8)a} & \text{for } k=3,\\ n^{(3/16)a} & \text{for } k=4,\\ 1 & \text{for } k\geqq 5.\end{cases}$$

By Lemma 18, we have

$$(6)\qquad \left| \sum_{\mathfrak{M}_1} \int_{\mathfrak{M}_1} \frac{f^s(x)}{x^{n+1}}\, dx - \sum_{\mathfrak{M}_1} \int_{\mathfrak{M}_1} \frac{\psi_\rho{}^s(x)}{x^{n+1}}\, dx \right| < C_{32} n^{sa-1-(s/K-2)a+\epsilon}$$
$$+ C_{33} n^{sa-1-a+\epsilon} \begin{cases} n^{(3/4)a} & \text{for } k = 3, \\ n^{(3/8)a} & \text{for } k = 4, \\ 1 & \text{for } k \geqq 5. \end{cases}$$

By Lemma 14,

$$(7)\qquad \left| \sum_{\mathfrak{M}_1} \int_{\mathfrak{M}_1} \frac{\psi_\rho{}^s(x)}{x^{n+1}}\, dx - \sum_{\mathfrak{M}_1} \int_{\mathfrak{K}} \frac{\psi_\rho{}^s(x)}{x^{n+1}}\, dx \right| < C_{14} n^{sa-1-(s/K-2)a+\epsilon}.$$

Further

$$\frac{1}{2\pi i} \sum_{\mathfrak{M}_1} \int_{\mathfrak{K}} \frac{\psi_\rho{}^s(x)}{x^{n+1}} = \alpha_0{}^{-sa} \frac{\Gamma(sa+n)}{n!} \frac{\Gamma^s(1+a)}{\Gamma(sa)} \sum_{\mathfrak{M}_1} (S_\rho/q^*)^s \rho^{-n}$$
$$= \alpha_0{}^{-sa} \frac{(sa+n)}{n!} \frac{\Gamma^s(1+a)}{\Gamma(sa)} \sum_{q=1}^{n^a} A_{P,s}(q,n).$$

Since by Lemma 7,

$$\left| \frac{\Gamma(sa+n)}{n!} - n^{sa-1} \right| < C_{46} n^{sa-2},$$

and by Lemma 10,

$$| A_{P,s}(q,n) | < C_{47} q^{1-s/K+\epsilon};$$

thus

$$(8)\qquad \left| \frac{1}{2\pi i} \sum_{\mathfrak{M}_1} \int_{\mathfrak{K}} \frac{\psi_\rho{}^s(x)}{x^{n+1}}\, dx - \alpha_0{}^{-sa} n^{sa-1} \frac{\Gamma^s(1+a)}{\Gamma(sa)} \sum_{q=1}^{n^a} A_{P,s}(q,n) \right|$$
$$< C_{48} n^{sa-2} \sum_{q=1}^{n^a} q^{1-s/K+\epsilon} < C_{48} n^{sa-2}.$$

Since

$$\left| \sum_{q>n^a}^{\infty} A_{P,s}(q,n) \right| < C_{49} \sum_{q>n^a}^{\infty} q^{1-s/K+\epsilon} < C_{50} n^{-(s/K-2)a+\epsilon},$$

we have

$$\left| \alpha_0{}^{-sa} n^{sa-1} \frac{\Gamma^s(1+a)}{\Gamma(sa)} \sum_{q=1}^{n^a} A_{P,s}(q,n) - \alpha_0{}^{-sa} n^{sa-1} \frac{\Gamma^s(1+a)}{\Gamma(sa)} \mathfrak{S}_n \right|$$
$$< C_{51} n^{-(s/K-2)a+sa-1+\epsilon}.$$

Then from (4), (5), (6), (7), and (8) the desired result follows.

TSING HUA UNIVERSITY (CHINA).

LIST OF REFERENCES.

1. E. Landau, "Zum Waringschen Problem," *Mathematische Zeitschrift*, Bd. 12 (1922), pp. 219-247.

2. E. Landau, "Über die neue Winogradoffsche Behandlung des Waringschen Problems," *Mathematische Zeitschrift*, Bd. 31 (1929), pp. 318-338.

3. M. Gelbcke, "Zum Waringschen Problem," *Mathematische Annalen*, Bd. 195 (1931), pp. 637-752.

4. R. D. James, "The representation of integers as sums of values of cubic polynomials," *American Journal of Mathematics*, vol. 56 (1934), pp. 303-315.

5. L. K. Hua, "On Waring theorems with cubic polynomial summands," *Mathematische Annalen*, Bd. 111, pp. 622-628.

6. E. Landau, *Vorlesungen über Zahlentheorie.*

7. D. Hilbert, *Mathematische Annalen*, Bd. 36 (1890), p. 511.

POLYNOMIALS FOR THE n-ARY COMPOSITION OF NUMERICAL FUNCTIONS.

By D. H. Lehmer.

In discussing the general binary associative and commutative composition of numerical functions it has been essential to deal with a function $\psi(x_1, x_2)$ satisfying (for $n = 2$) the three postulates given below.[1] In this paper we find all polynomials $\psi(x_1, x_2, \cdots, x_n)$ satisfying these postulates and discuss some of their properties.

The postulates are as follows:

Postulate A. *If each x is a positive integer, $\psi(x_1, x_2, \cdots, x_n)$ is a positive integer.*

Postulate B. *For each positive integer N the equation $N = \psi(x_1, x_2, \cdots, x_n)$ has a solution in positive integers.*

Postulate C. *$\psi(x_1, x_2, \cdots, x_{n-1}, \psi(x_n, x_{n+1}, \cdots, x_{2n-1}))$ is a symmetric function.*

Postulate C is a generalization of the commutative and associative laws.[2] The function appearing in this postulate will for reasons of brevity be called the iterate of ψ, and will be designated by $\psi^* = \psi^*(x_1, x_2, \cdots, x_{2n-1})$. We begin by assuming Postulate C alone. From this point on ψ will be assumed to be a polynomial.[3]

Theorem 1. *If $\psi(x_1, x_2, \cdots, x_n)$ is a polynomial such that*

$$\psi^* = \psi(x_1, x_2, \cdots, x_{n-1}, \psi(y_1, y_2, \cdots, y_n))$$

---

[1] D. H. Lehmer, *Transactions of the American Mathematical Society*, vol. 33 (1931), pp. 945-957, *Bulletin of the American Mathematical Society*, vol. 37 (1931), pp. 723-726; G. Pall, *Bulletin of the American Mathematical Society*, vol. 38 (1932), pp. 56-58; E. T. Bell, *Bulletin of the American Mathematical Society*, vol. 41 (1935), pp. 798, 801.

[2] This postulate could have been weakened at the expense of simplicity in a manner analogous to that used in following papers: W. A. Hurwitz, *Annals of Mathematics* (2), vol. 15 (1913-14), p. 93; D. H. Lehmer, *American Journal of Mathematics*, vol. 54 (1932), p. 329.

[3] Abel, *Journal für Mathematik*, vol. 1, pp. 11-15; *Oeuvres*, vol. 1, pp. 61-66, discussed functions of two variables satisfying Postulate C and stated that such functions must be symmetric. Examples to the contrary can be given however. For instance let $\psi(1, 2) = 3$, while $\psi(x, y) = 1$ otherwise. Here $\psi(x, \psi(y, z))$ is symmetric whereas $\psi(1, 2) \neq \psi(2, 1)$.

Lémeray, *Nouvelles Annales Mathématiques* (4), vol. 1 (1901), pp. 163-167 used symmetric functions satisfying Postulate C to generalize the Dedekind inversion formula.

is a symmetric function of its $2n-1$ variables, then ψ is symmetric and linear in each of its n variables.

Proof. Let $\psi(x_1, x_2, \cdots, x_n)$ be of degree $d_k > 0$ in x_k $(k = 1, 2, \cdots, n)$. Then ψ^* is of degree d_1 in x_1 and of degree $d_k d_n$ in y_k. But it is symmetric by hypothesis. Hence $d_1 = d_k d_n$. Since $d_1 \neq 0$, it follows with $k = 1$, that $d_n = 1$. Hence $d_1 = d_2 = \cdots = d_n = 1$, so that ψ is linear in each variable. The symmetry of ψ now follows from the symmetry of ψ^* since ψ^* is linear in ψ.

If we write

$$\sigma_{n,k} = \sigma_{n,k}(x_1, x_2, \cdots, x_n) = \Sigma x_1 x_2 \cdots x_k$$

for the k-th elementary symmetric function of $x_1, x_2, \cdots, x_n$ we may (by Theorem 1) write our ψ function in the form

$$\psi(x_1, x_2, \cdots, x_n) = a_0 + a_1\sigma_{n,1} + \cdots + a_n\sigma_{n,n}. \tag{2}$$

Our problem is to find necessary and sufficient conditions on the a's so that ψ^* will be symmetric.

If in (2) we collect coefficients of all products containing x_n we get

$$\psi(x_1, x_2, \cdots, x_n) = \sum_{\nu=0}^{n-1} (a_{\nu+1}x_n + a_\nu)\sigma_{n-1,\nu}(x_1, x_2, \cdots, x_{n-1}).$$

Hence

$$\begin{aligned} \psi^* &= \psi(x_1, x_2, \cdots, x_{n-1}, \psi(y_1, y_2, \cdots, y_n)) \\ &= \sum_{\nu=0}^{n-1} (a_{\nu+1}\psi(y_1, y_2, \cdots, y_n) + a_\nu)\sigma_{n-1,\nu}(x_1, x_2, \cdots, x_{n-1}). \end{aligned} \tag{3}$$

Now the coefficient of any product occurring in (3) and consisting of precisely k x's and $r-k$, y's $(0 \leqq k \leqq n-1;\ r-k \leqq n)$ is given by

$$\begin{array}{ll} a_{k+1}a_{r-k} & \text{if} \quad 0 \leqq k < r \\ a_r + a_{r+1}a_0 & \text{if} \quad 0 \leqq k = r. \end{array} \tag{4}$$

For ψ^* to be symmetric it is necessary that for each r and for different values of $k \leqq r$, the coefficients (4) should be equal. Two cases now present themselves according as $a_0 = 0$ or not.

Case 1. $a_0 \neq 0$. Let $R = (a_1 - 1)/a_0$. Let $r = 1$, and $k = 0, 1$ in (4) and equate. We have

$$a_1{}^2 = a_1 + a_2 a_0.$$

That is

$$a_2 = a_1 R.$$

By induction, if $a_i = a_1 R^{i-1}$, for some $i < n$, we can determine a_{i+1} by setting $r = i$, and $k = i - 1$ and i in (4) and equating. We have

$$a_1 a_i = a_i + a_{i+1} a_0,$$

whence

$$a_{i+1} = R\, a_i = a_1 R^i.$$

Hence in general

(5) $$a_\nu = a_1 R^{\nu-1} \qquad (\nu = 1, 2, \cdots, n).$$

Case II. $a_0 = 0$. In this case the coefficients (4) become

(4′) $$\begin{array}{lll} a_{k+1} a_{r-k} & \text{if} & 0 \leqq k < r \\ a_r & \text{if} & k = r. \end{array}$$

Setting $r = 1$, and $k = 0, 1$ we get on equating

$$a_1^2 = a_1.$$

Hence two subcases arise according as $a_1 = 0$ or 1.

Case II$_1$. $a_0 = 0$, $a_1 = 1$. In this case no condition on a_2 is obtained. To determine a_3 we set $r = 3$ and $k = 0, 1$ in (4′) and get

$$a_3 = a_2^2.$$

By induction we find that

$$a_\nu = a_2^{\nu-1}, \qquad 2 \leqq \nu \leqq n, \qquad a_0 = 0, \quad a_1 = 1.$$

Case II$_2$. $a_0 = a_1 = 0$. Setting $k = r - 1$ and r in (4′) with $r < n$, we find

$$a_r a_1 = a_r = 0.$$

For $r = n$, however, we get no condition on a_n. Hence in this case

$$\psi(x_1, x_2, \cdots, x_n) = a_n x_1 x_2 \cdots x_n.$$

We may sum up the results obtained thus far in the following

THEOREM 2. *A polynomial* $\psi(x_1, x_2, \cdots, x_n)$, *whose iterate*

$$\psi^*(x_1, x_2, \cdots, x_{2n-1}) = \psi(x_1, x_2, \cdots, x_{n-1}, \psi(x_n, \cdots, x_{2n-1}))$$

is symmetric, is necessarily of one of the forms

(6) $$\psi(x_1, x_2, \cdots, x_n) = a_0 + a_1 \Sigma x_i + \frac{a_1(a_1 - 1)}{a_0} \Sigma x_i x_j + \frac{a_1(a_1-1)^2}{a_0^2} \Sigma x_i x_j x_k + \cdots + \frac{a_1(a_1-1)^{n-1}}{a_0^{n-1}} x_1 x_2 \cdots x_n.$$

(7) $$\psi(x_1, x_2, \cdots, x_n) = \Sigma x_i + a_2 \Sigma x_i x_j + a_2^2 \Sigma x_i x_j x_k + \cdots + a_2^{n-1} x_1 \cdots x_n.$$

(8) $$\psi(x_1, x_2, \cdots, x_n) = a_n x_1 x_2 \cdots x_n.$$

We now show that all polynomials of the form (6), (7) or (8) have symmetric iterates. In fact this is obvious in case ψ is of the form (8). The same is true of (7) when $a_2 = 0$. If $a_2 \neq 0$, we can write (7) in the form

$$a_2\psi(x_1, x_2, \cdots, x_n) + 1 = \prod_{i=1}^{n} (a_2x_i + 1). \tag{7'}$$

Hence

$$\begin{aligned} a_2\psi^* + 1 &= a_2\psi(x_1, x_2, \cdots, x_{n-1}, \psi(x_n, x_{n+1}, \cdots, x_{2n-1})) + 1 \\ &= \prod_{i=1}^{n-1} (a_2x_i + 1)\{a_2\psi(x_n, \cdots, x_{2n-1}) + 1\} \\ &= \prod_{i=1}^{2n-1} (a_2x_i + 1). \end{aligned} \tag{9}$$

Therefore ψ^* is symmetric. If ψ is of the form (6) and $a_1(a_1 - 1) \neq 0$ (the symmetry of ψ^* being evident otherwise), we may write (6) in the form

$$R\psi(x_1, x_2, \cdots, x_n) + 1 = a_1 \prod_{i=1}^{n} (Rx_i + 1), \quad R = (a_1 - 1)/a_0 \neq 0. \tag{6'}$$

As before we find

$$R\psi^* + 1 = a_1^2 \prod_{i=1}^{2n-1} (Rx_i + 1) \tag{10}$$

so that ψ^* is symmetric in this case also. *Hence the class of all polynomials whose iterates are symmetric coincides with* (6), (7) *and* (8).

Converse propositions regarding product representation may also be proved.

THEOREM 3. *Let $\psi(x_1, x_2, \cdots, x_n)$ be a symmetric multilinear form*

$$a_0 + a_1\Sigma x_i + a_2\Sigma x_ix_j + \cdots + a_nx_1x_2\cdots x_n$$

in which $a_0a_1(a_1 - 1) \neq 0$. Suppose further that $1 + \psi(a_1 - 1)/a_0$ be expressible as a product of linear factors. Then the iterate of ψ is also symmetric and ψ is of type (6).

Proof. Let $R = (a_1 - 1)/a_0$. By hypothesis α_i and β_i exist such that

$$R\psi(x_1, x_2, \cdots, x_n) + 1 = \prod_{i=1}^{n} (\alpha_ix_i + \beta_i).$$

If we set all the variables equal to zero, we obtain

$$Ra_0 + 1 = \prod_{i=1}^{n} \beta_i = a_1 \neq 0.$$

Hence we can set $\alpha_i/\beta_i = \gamma_i$ and write

$$R\psi(x_1, x_2, \cdots, x_n) + 1 = a_1 \prod_{i=1}^{n} (\gamma_ix_i + 1).$$

But ψ is a symmetric function so that the γ's are equal. Hence

$$(11) \qquad R\psi + 1 = a_1 \prod_{i=1}^{n} (\gamma x_i + 1).$$

Setting $x_1 = 1$ and $x_i = 0$ for $i > 1$, we have

$$(12) \qquad Ra_0 + Ra_1 + 1 = a_1(\gamma + 1).$$

Recalling the definition of R we obtain from (12)

$$a_1(R + 1) = a_1(\gamma + 1).$$

Hence $R = \gamma$, and (11) becomes (6′). Therefore ψ^* is symmetric and ψ is of type (6).

THEOREM 4. *Let $\psi(x_1, x_2, \cdots, x_n)$ be of the form*

$$a_0 + \Sigma x_i + a_2 \Sigma x_i x_j + \cdots + a_n x_1 x_2 \cdots x_n$$

where $a_2 \neq 0$. Suppose further that $a_2\psi + 1$ is expressible as a product of linear factors. Then ψ^* is symmetric and ψ is of the type (7).

Proof. By hypothesis we have

$$a_2\psi(x_1, x_2, \cdots, x_n) + 1 = \prod_{i=1}^{n} (\alpha_i x_i + \beta_i).$$

Setting all the variables equal to zero we have

$$a_2 a_0 + 1 = \prod_{i=1}^{n} \beta_i.$$

We show now that each β is different from zero. In fact if one β were zero, all would have to be zero since $a_2\psi(x_1, x_2, \cdots, x_n) + 1$ is a symmetric function. Moreover we would have

$$(13) \qquad a_0 a_2 + 1 = 0$$

$$(14) \qquad a_2\psi(x_1, x_2, \cdots, x_n) + 1 = \prod_{i=1}^{n} \alpha_i x_i.$$

Setting $x_1 = 1$, and $x_i = 0$ for $i > 1$ in (14) we would have

$$a_2 a_0 + a_2 + 1 = 0.$$

This contradicts (13). Hence $\Pi\beta_i = B \neq 0$, and we can write

$$(15) \qquad a_2\psi(x_1, x_2, \cdots, x_n) + 1 = B \prod_{i=1}^{n} (\gamma_i x_i + 1)$$

where $\gamma_i = \alpha_i/\beta_i$. By symmetry the γ's are equal.

Expanding (15) with this fact in mind we have

$$1 + a_2 a_0 + a_2 \sigma_{n,1} + a_2{}^2 \sigma_{n,2} + \cdots = B + B\gamma\sigma_{n,1} + B\gamma^2\sigma_{n,2} + \cdots.$$

Identifying coefficients of the σ's we have

$$B = 1 + a_2 a_0, \qquad B\gamma = a_2, \qquad B\gamma^2 = a_2{}^2.$$

Hence, $B = 1$, $a_0 = 0$, $a_2 = \gamma$. Therefore ψ is of the type (7′) and hence of type (7). The symmetry of ψ^* follows as before.

THEOREM 5. *If the iterate ψ^* of a polynomial ψ is symmetric so also is the iterate of ψ^*.*

Proof. In case ψ is a simple sum or product, the theorem is obvious. If not, ψ will be of type (6) or (7). By (9) and (10) ψ^* will be of type (6) or (7) also. Hence the iterate of ψ^* will be symmetric.

THEOREM 6. *Every polynomial ψ in n variables, whose iterate ψ^* is symmetric may be obtained by iterating $n - 2$ times such a polynomial in 2 variables.*

Proof. In case ψ is a simple sum or product the theorem is obvious. Otherwise ψ will be of the type (6) or (7) with $a_1 \neq 1$. In the first case ψ is the result of $n - 2$ successive iterations of the function

$$\psi(x_1, x_2) = A_0 + A_1(x_1 + x_2) + \frac{A_1(A_1 - 1)}{A_0} x_1 x_2$$

where

$$A_1{}^{n-1} = a_1 \tag{16}$$

$$A_0 = a_0(A_1 - 1)/(a_1 - 1). \tag{17}$$

In fact the result of iterating $\psi(x_1, x_2)$ a certain number of times is in every case a function of the type (6), the ratio of consecutive coefficients being always $(A_1 - 1)/A_0$ by (9). Hence (17) must hold. The coefficients of x_1 in the successive iterates of $\psi(x_1, x_2)$ are the successive powers of A_1. Hence (16). In case ψ is of type (7), it is seen to be the result of iterating

$$\psi(x_1, x_2) = x_1 + x_2 + a_2 x_1 x_2.$$

Hence the theorem.

We now turn our attention to Postulates A and B. To simplify the discussion we introduce a function f by

(18) $f(x_1, x_2, \cdots, x_n) = -1 + \psi((x_1+1), (x_2+1), \cdots, (x_n+1))$.

By Theorem 1, f will be a symmetric multilinear form

(19) $f(x_1, x_2, \cdots, x_n) = c_0 + c_1\Sigma x_i + c_2\Sigma x_i x_j + \cdots + c_n x_1 x_2 \cdots x_n$.

If Postulates A and B hold for ψ, f must, by (18), satisfy the conditions.

(A$'$). *If each x is an integer $\geqq 0$, so also is $f(x_1, x_2, \cdots, x_n)$.*

(B$'$). *The equation $f(x_1, x_2, \cdots, x_n) = N$ has a solution $(x_1, x_2, \cdots, x_n)$ in integers $\geqq 0$ for each $N \geqq 0$.*

THEOREM 7. *A function f of the form* (19) *satisfies* (A$'$) *and* (B$'$) *if and only if $c_0 = 0$, $c_1 = 1$ while $c_2, c_3, \cdots, c_n$ are integers $\geqq 0$.*

Proof. The sufficiency of the condition as far as (A$'$) is concerned is obvious, and (B$'$) is easily seen to follow from the identity

$$f(N, 0, 0, \cdots, 0) = N.$$

As to necessity, we first show that if f satisfies (A$'$) then the c_ν are integers $\geqq 0$. In fact if we set $x_1 = x_2 = \cdots = x_\nu = x$ and $x_{\nu+1} = x_{\nu+2} = \cdots = x_n = 0$, we obtain from (19)

$$f = c_0 + \binom{\nu}{1} c_1 x + \binom{\nu}{2} c_2 x^2 + \cdots + c_\nu x^\nu. \tag{20}$$

If c_k is an integer for all $k < \nu$, it follows by setting $x = 1$ that c_ν is an integer in view of (A$'$). But $c_0 = f(0, 0, \cdots, 0)$ is an integer. Hence by induction each c is an integer. However if c_ν were negative, we could choose x in (20) so large that f would be negative contrary to (A$'$). Hence the c's are non-negative integers.

To show that $c_0 = 0$ and $c_1 = 1$ we introduce (B$'$) and note first that not all the c's are zero since $f =$ constant fails to satisfy (B$'$). For this reason

$$f(x_1, x_2, \cdots, x_n) > f(0, 0, \cdots, 0) = c_0 \geqq 0,$$

provided the x's are non-negative integers not all zero. By (B$'$) f must represent zero. Hence $c_0 = 0$. Similarly the next largest value of f is

$$f(1, 0, 0, \cdots, 0) = c_0 + c_1 = c_1.$$

Hence $c_1 = 1$ and the theorem follows. Another statement of it is

THEOREM 8. *All symmetric multilinear forms ψ satisfying Postulates* A *and* B *are*

$$(21)\qquad \psi(x_1, x_2, \cdots, x_n) = 1 + \Sigma(x_i - 1) + c_2\Sigma(x_i - 1)(x_j - 1) + \cdots + c_n(x_1 - 1)(x_2 - 1)\cdots(x_n - 1)$$

where the c's are not-negative integers.

Writing ψ in the form

$$\psi(x_1, x_2, \cdots, x_n) = a_0 + a_1\Sigma x_i + a_2\Sigma x_i x_j + \cdots + a_n x_1 x_2 \cdots x_n$$

and identifying the coefficients with those obtained after expanding the right side of (21), we obtain the relation

$$(22)\qquad a_\nu = \sum_{\lambda=0}^{n-\nu} c_{\nu+\lambda}(-1)^\lambda \binom{n-\nu}{\lambda}$$

or what is the same

$$(23)\qquad c_\nu = \sum_{\lambda=0}^{n-\nu} a_{\nu+\lambda} \binom{n-\nu}{\lambda}.$$

It follows from (23) by a simple induction that the a's are integers.

Thus far we have not made full use of Postulate C. To do this we require that ψ be of the form (6), (7) or (8).

In the first of these three cases $a_k = R^{k-1}a_1$, $(k = 1, 2, \cdots, n)$. Putting these values in (23) we have for $\nu = 1$

$$c_1 = 1 = \sum_{\lambda=0}^{n-1} a_1 R^\lambda \binom{n-1}{\lambda} = a_1(1 + R)^{n-1}.$$

Recalling that $R = (a_1 - 1)/a_0$, we have

$$(24)\qquad a_0^{n-1} = a_1(a_0 + a_1 - 1)^{n-1}.$$

Since the a's are integers it follows that a_1 is the $(n-1)$-st power of an integer, say $a_1 = a^{n-1}$. Substituting this in (24) and taking $(n-1)$-st roots we find for every n

$$(25)\qquad a_0 = (a^n - a)/(1 - a), \quad R = (1 - a)/a, \quad a_\nu = (1 - a)^{\nu-1}a^{n-\nu}.$$

In case $n - 1$ is even we obtain another solution of (24)

$$(26)\qquad a_0 = -(a^n - a)/(1 + a), \qquad R = -(1 + a)/a,$$
$$a_\nu = (1 + a)^{\nu-1}a^{n-\nu}(-1)^{\nu-1}.$$

Substituting these values of a_ν in (23) we have for $\nu > 1$

$$c_\nu = (1 \pm a)^{\nu-1}$$

according as (26) or (25) is taken. Since $c_\nu \geqq 0$ we find in these respective cases

$$a \geqq 1 \quad \text{or} \quad a \leqq 1.$$

The polynomials found thus far are therefore

$$(27)\quad \psi(x_1, x_2, \cdots, x_n) = (a^n - a)/(1-a) + a^{n-1}\Sigma x_i + (1-a)a^{n-2}\Sigma x_i x_j + (1-a)^2 a^{n-3}\Sigma x_i x_j x_k + \cdots + (1-a)^{n-1} x_1 x_2 \cdots x_n$$

with a an integer $\leqq 1$, and

$$(28)\quad \psi(x_1, x_2, \cdots, x_n) = -(a^n - a)/(1+a) + a^{n-1}\Sigma x_i - (1+a)a^{n-2}\Sigma x_i x_j + (1+a)^2 a^{n-3}\Sigma x_i x_j x_k + \cdots + (1+a)^{n-1} x_1 x_2 \cdots x_n$$

where n is odd and a is an integer $\geqq 1$.

Returning now to polynomials of type (7) and (8) we find that those which satisfy Postulates A and B are of the type (27) or (28).

In fact if ψ is of type (7) we may substitute $a_k = a_2^{k-1}$ $(k = 1, 2, \cdots, n)$ in (23) and with $\nu = 1$ obtain

$$(29)\qquad c_1 = 1 = \sum_{\lambda=0}^{n-1} a_2^\lambda \binom{n-1}{\lambda} = (1 + a_2)^{n-1}$$

If n is even (29) has the single solution $a_2 = 0$, which leads to the polynomial

$$\psi(x_1, x_2, \cdots, x_n) = x_1 + x_2 + \cdots + x_n$$

which fails to satisfy Postulate B with $N = 1$. If n is odd (29) has the additional solution $a_2 = -2$. This leads to the same polynomial as the special case of (27) in which $a = -1$.

If ψ is of the type (8) namely

$$\psi(x_1, x_2, \cdots, x_n) = a_n x_1 x_2 \cdots x_n,$$

it is clear that Postulate B will be satisfied if and only if $a_n = 1$. But this is (27) with $a = 0$. Finally we observe that (28) is a duplication of (27) since every form obtained from (28) with $a \geqq 1$ is identical with the form obtained from (27) with $-a$.

Hence we may sum up the investigation by the following

THEOREM 9. *All polynomials which satisfy Postulates* A, B *and* C *are given by* (27), *where a is an integer* $\leqq 1$.

For $n = 2$, (27) becomes [4]

$$\psi(x_1, x_2) = -a + a(x_1 + x_2) + (1 - a)x_1 x_2.$$

By (16) and (17) this polynomial if iterated $n - 2$ times produces (27). Hence

THEOREM 10. *There exist no polynomials satisfying Postulates* A, B *and* C *which are not iterates of a polynomial in two variables, satisfying these postulates.*

The effect of Theorem 10 on the theory of n-ary composition of numerical functions may be described in a few lines.

Let $f_\nu(N)$ $(\nu = 1, 2, \cdots, n)$ be n arbitrary numerical functions defined for all positive integers N. The composite (or symbolic product) of these functions with respect to a function $\psi(x_1, x_2, \cdots, x_n)$ is a function $F(N)$ which may be defined for example by

$$F(N) = \Sigma f_1(x_1) f_2(x_2) \cdots f_n(x_n)$$

where the sum extends over all solutions $(x_1, x_2, \cdots, x_n)$ of $\psi(x_1, x_2, \cdots, x_n) = N$ and where the multiplication indicated under the sign of summation may be ordinary multiplication for example. If ψ satisfies Postulates A, B and C, then $F(N)$ will be a numerical function with the maximum degree of symmetry. As a result of Theorem 10 any composition of numerical functions n at a time with respect to a polynomial can be achieved by repeated applications of the ordinary binary composition.

LEHIGH UNIVERSITY,
BETHLEHEM, PA.

[4] This polynomial is identical with that of Pall, *loc. cit.*, with $a - 1 = \alpha$.

MINIMUM PARTITIONS INTO SPECIFIED PARTS.

By Hansraj Gupta.

1. Dickson[1] has studied "Minimum Decompositions" of numbers into "N-th powers." In this paper, I consider the problem of partitions of a positive integral number n into parts $1, a, b$; where a, b are positive integers such that

$$(1) \qquad 1 < a < b.$$

The problem is here attacked directly and in an elementary way.

2. If

$$(2) \qquad n = x + ay + bz,$$

where x, y, z are positive integers ≥ 0; then $(x + y + z)$ shall be called in general the weight of n. The least value of $(x + y + z)$ shall be termed after Dickson the "Minimum weight" of n, and written Min. (n).

If

$$(3) \qquad n = X + aY + bZ, \quad \text{where} \quad 0 \leq X < a,\ 0 \leq X + aY < b;$$

then $(X + Y + Z)$ shall be termed the "Absolute weight" of n, and written Ab. (n). Evidently

$$(4) \qquad n \geq \text{Ab. } (n) \geq \text{Min. } (n).$$

3. Let

$$(5) \qquad b = qa + r, \qquad 0 \leq r < a.$$

If $r = 0$, then Min. $(n) =$ Ab. (n). If

$$(6) \qquad r > 0, \quad \text{let} \quad X + jr = ia + m, \quad 0 \leq m < a, \ 0 \leq j \leq Z;$$

then

$$(7) \qquad n = m + a(Y + jq + i) + b(Z - j).$$

Now for some value of j, (7) must represent the partition with minimum weight. Subtracting the weight of (7) from the absolute weight, we get

$$(8.1) \qquad \Delta \equiv X - (m + i) - j(q - 1),$$

or

$$(8.2) \qquad \Delta(i, j) \equiv i(a - 1) - j(q + r - 1).$$

[1] *American Journal of Mathematics*, vol. 55 (1933).

If Min. $(n) \neq$ Ab. (n), we must have

(9) $$1 \leq \frac{j}{i} < \frac{a-1}{q+r-1}.$$

Hence

(10) Min. $(n) =$ Ab. (n), if $q+r \geq a$; in particular if $b \geq a.^2$

Moreover as j takes the values $1, 2, 3, \cdots$ in (6), m decreases only when i increases. Hence to find the maximum value of Δ, we need give the values $1, 2, 3, \cdots$ to i only.

For any value of i, j will be the least if $0 \leq m < r$. Hence

(11) $$j = \left[\frac{ia - X + r - 1}{r}\right],$$

with the condition stated in (6).

We can now express $\Delta(i, j)$ as a function of i alone. Thus

(12) $$\Delta(i) = i(a-1) - \left[\frac{ia - X + r - 1}{r}\right](q+r-1).$$

4. $\Delta(i)$. In (12) if i is increased by t, then we have either

(13. 1) $$\delta_1(t) \equiv \Delta(i+t) - \Delta(i) = t(a-1) - \left[\frac{ta}{r}\right](q+r-1);$$

or

(13. 2) $$\delta_2(t) \equiv \Delta(i+t) - \Delta(i) = t(a-1) - \left[\frac{ta+r}{r}\right](q+r-1).$$

The latter is always negative.

If $s > 0$ be a value of t, for which $\delta_1(t)$ is negative, then

(14) $$\Delta(i+s) < \Delta(i).$$

This result is independent of X.

Hence to find the maximum value of $\Delta(i)$, i need not be given values $> s$ in (12).

5. Search for s. Change a/r and $(a-1)/(q+r-1)$ into simple continued fractions, and find an odd convergent c_{2l+1} of the former $\geq$ an even convergent C_{2k} of the latter. If d_h denote the denominator of the h-th convergent of a/r, then a value of s will be found among the members of the A. P. whose first term is $(d_{2l} + d_{2l-1})$ and common difference is d_{2l}.

An example will make the method clear. Let

$$a = 2^{22} = 4194304, \qquad b = 3^{22};$$

then

$$q = 7481, \quad \text{and} \quad r = 3471385.$$

Changing a/r and $(a-1)/(q+r-1)$ into s. c. fractions, we get the convergents

$$\text{(c)} \qquad \frac{1}{1}, \frac{5}{4}, \frac{6}{5}, \frac{29}{24}, \frac{586}{485}, \cdots$$

and

$$\text{(C)} \qquad \frac{1}{1}, \frac{5}{4}, \frac{6}{5}, \frac{41}{34}, \cdots.$$

We find

$$\frac{586}{485} > \frac{41}{34}.$$

We now search for s among the numbers 29, 53, 77, 101, $\cdots$; and find $\delta_1(29) < 0$. Hence $s = 29$. We notice that $\delta_1(t) \nless 0$ for any value of $t < 29$.

6. The following table gives s, when

$$a = 2^u, \quad b = 3^u; \qquad u \leq 36.$$

The values of q and r, for these values of u are given in Dickson's paper cited above.

u	1	2	3	4	5	6	7	8	9	10
s	1	1	1	1	2	1	1	2	1	2
u	11	12	13	14	15	16	17	18	19	20
s	1	3	5	14	9	16	5	17	21	7
u	21	22	23	24	25	26	27	28	29	30
s	31	29	23	7	12	64	4	70	4	6
u	31	32	33	34	35	36				
s	53	53	245	73	20	9				

7. The numbers X from 0 to $(a-1)$ can be divided into a number of groups; all numbers for which Δ has the same maximum value being placed in one group. Thus X will be said to belong to the group (c, e), if for the

given X, $\Delta(i,j)$ is maximum when $i = c$, and $j = e$. The inferior limits of these groups are easily obtained. Thus the inferior limit of (c, e) is

$$ac - re. \tag{15}$$

Since $ac - re < a$, therefore $e > a(c-1)/r$. Moreover

$$e < (a-1)c/(q+r-1).$$

Hence

$$\left[\frac{a(c-1)}{r}\right] < e \leq \left[\frac{(a-1)c}{q+r-1}\right]. \tag{16}$$

As an example, consider $a = 2^{13} = 8192$; $b = 3^{13} = 1594323$; then $q = 194$, $r = 5075$; and $s = 5$. The limits for the various groups are tabulated below.

Group (c, e)	Inferior Limit $ac - re$	Superior Limit	Maximum Δ $(a-1)c - (q+r-1)e$
(0, 0)	0	1158	0
(1, 1)	3117	4275	2923
(2, 2)	6234	7392	5846
(2, 3)	1159	2317	578
(3, 4)	4276	5434	3501
(4, 5)	7393	8191	6424
(4, 6)	2318	3116	1156
(5, 7)	5435	6233	4079

The superior limits are ascertained after the inferior limits have been calculated.

GOVERNMENT COLLEGE,
HOSHIARPUR, INDIA.

DIVISIBILITY SEQUENCES OF THIRD ORDER.

By Marshall Hall.

1. *Introduction.* By a divisibility sequence of k-th order will be meant a sequence of rational integers $u_0, u_1, u_2, \cdots u_n, \cdots$ satisfying the linear recurrence

$$u_{n+k} = a_1 u_{n+k-1} + \cdots + a_k u_n \tag{1}$$

where the a's are rational integers, and such that $u_n | u_{mn}$ (read u_n divides u_{mn}) for any m and n not zero.

It will be shown (a) that there are two types of divisibility sequences which may be distinguished according as $u_0 \neq 0$ or $u_0 = 0$. If $u_0 \neq 0$, the totality of primes dividing terms of the sequence is finite and the sequence is said to be degenerate. If $u_0 = 0$, all but a finite number of primes will appear as divisors of the terms, and we call the sequence regular. Furthermore this paper shows (b) that the factorization properties of divisibility sequences are similar to the factorization properties of the Lucas[1] sequences, and (c) that there is no regular divisibility sequence of third order whose associated cubic is irreducible. Here a_2 and a_3 are assumed to be co-prime.

Divisibility sequences are of particular interest because of their remarkable factorization properties. Lucas was the first to discover the striking relations in second order sequences and give a coherent theory, though some of his results were implied by earlier work on the theory of quadratic forms. Among other results, he developed the tests for primality applicable to the Mersenne numbers. Other special types of divisibility sequences have been investigated by Lehmer,[2] Pierce,[3] and Poulet.[4]

2. *Properties of General Linear Recurrences.* There will be occasion to use the following properties of recurring sequences, whether divisibility sequences or not. Let the sequence (u_n) be determined by the recurrence

[1] E. Lucas, "Théorie des Fonctions Numériques Simplement Périodiques," *American Journal of Mathematics*, vol. 1 (1875), pp. 184-240, 289-321.

[2] D. H. Lehmer, "An extended theory of Lucas' functions," *Annals of Mathematics* (2), vol. 31 (1930), pp. 419-448.

[3] T. A. Pierce, "The numerical factors of the arithmetic forms $\prod_{i=1}^{n} (1 \pm \alpha_i^m)$," *Annals of Mathematics* (2), vol. 18 (1916-17), pp. 53-64.

[4] Poulet, *L'Intermédiare des Mathématiciens*, vol. 27, pp. 86-87; (2), vol. 1, p. 47; vol. 3, p. 61.

(1) and by an initial set of values $u_0, u_1, u_2, \cdots u_{k-1}$. With the recurrence (1) is associated its characteristic polynomial,

$$f(x) = x^k - a_1 x^{k-1} \cdots - a_k = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_k).$$

If the roots of $f(x)$ are distinct, then

$$u_n = c_1 \alpha_1{}^n + c_2 \alpha_2{}^n \cdots + c_k \alpha_k{}^n \tag{2}$$

where $c_1, c_2, \cdots c_k$ are constants which may be determined from the initial values $u_0, u_1, \cdots u_{k-1}$.

The sequence (u_n) is periodic[5] for an arbitrary modulus m. That is to say there exists a period τ of (u_n) modulo m, depending on m and $a_1, a_2, \cdots a_k$ such that

$$u_{n+\tau} \equiv u_n \pmod m \tag{3}$$

for all $n \geqq n_0[m, a_1, a_2, \cdots a_k]$. In particular $n_0 = 0$ if $(a_k, m) = 1$. The period τ is taken to be the least number satisfying such a relation. All other numbers with this property are multiples of the period. If p be a prime not dividing the discriminant of $f(x)$, and if $f(x) \equiv f_1(x) f_2(x) \cdots f_s(x) \pmod p$ be the decomposition of $f(x)$ into irreducible factors modulo p, whose degrees are $k_1, k_2, \cdots k_s$ respectively, then τ divides the least common multiple of $p^{k_i} - 1$, $i = 1, 2, \cdots s$. Moreover (u_n) has a restricted period[6] $\mu \pmod m$. μ is defined to be the least integer for which there is a b such that

$$u_{n+\mu} \equiv b u_n \pmod m \tag{4}$$

for all $n \geqq n_0$. If e is the exponent to which b belongs $\pmod m$, then $\mu e = \tau$. If $f(x)$ is irreducible modulo p, p a prime, then $\mu \mid \dfrac{p^k - 1}{p - 1}$.

3. *Properties of General Divisibility Sequences.* References have been given above to investigations of certain types of divisibility sequences. This paper, however, is the first to treat them in general. It is the first attempt to find what the general characteristics of a divisibility sequence are, and what types exist. In this section the fundamental difference between regular and degenerate divisibility sequences is given by Theorem II. Theorem III is the key to the factorization properties of all divisibility sequences. In § 4 these theorems are applied to third order divisibility sequences.

---

[5] H. T. Engstrom, "On sequences defined by linear recurrence relations," *Transactions of the American Mathematical Society*, vol. 33 (1931), pp. 210-218.

[6] R. Carmichael, "On sequences of integers defined by recurrence relations," *Quarterly Journal of Mathematics*, vol. 48 (1920), pp. 343-372. See page 354 for reference to the restricted period. In particular $(b, m) = 1$ if $(a_k, m) = 1$.

THEOREM I. *If (u_n) is a divisibility sequence and some u_r has a factor m relatively prime to a_k, then $u_0 \equiv 0 \pmod{m}$.*

As (u_n) is a divisibility sequence $u_r | u_{\tau r}$, and hence $u_{\tau r} \equiv 0 \pmod{m}$. Since $(a_k, m) = 1$, relation (3) holds with $n = 0$. This yields $u_{\tau r} \equiv u_0 \pmod{m}$ and hence $u_0 \equiv 0 \pmod{m}$ as was to be proved.

It is on the basis of this theorem that divisibility sequences have been separated into two categories, viz., degenerate if $u_0 \neq 0$, regular if $u_0 = 0$.

If u_n be any term of a degenerate divisibility sequence (u_n), it may be written as the product of two factors, $u_n = A_n B_n$, where $A_n | u_0$, and B_n is divisible only by primes dividing a_k. The totality of primes dividing the terms of (u_n) will be finite. Degenerate divisibility sequences will be excluded from consideration in this paper, but will be treated further elsewhere.

If (u_n) is a regular divisibility sequence satisfying (1) and p is any prime not dividing a_k, $u_{s\tau} \equiv u_0 \equiv 0 \pmod{p}$ where τ is the period of (u_n) modulo p. Hence every prime not dividing a_k will divide the terms of a subsequence of (u_n) if (u_n) is a regular divisibility sequence. Furthermore, we may take $u_1 = 1$ without loss of generality since $(u_n) = (v_n/v_1)$ is a divisibility sequence satisfying (1) if (v_n) is a divisibility sequence satisfying (1). (u_n) will, of course, be a sequence of integers as $v_1 | v_n$ for all n, including $n = 0$, as $v_0 = 0$.

It is convenient to state these results as a theorem.

THEOREM II. *The totality of primes dividing the terms of a degenerate sequence (u_n) is contained in the set of primes dividing u_0 and a_k. The totality of primes dividing the terms of a regular sequence (u_n) includes every prime not dividing a_k.*

Consider the factorization of u_n, a particular term of a regular divisibility sequence. By the divisibility property, any prime dividing u_r where $r | n$ is a divisor of u_n. The remaining primes belong essentially to the term u_n itself.

Definition. A prime p is said to be a primitive divisor of u_n if $p | u_n$, $p \nmid u_r$ for $r | n$, $r \neq n$, and if $p \nmid a_k$.

The following theorem on the factorization of terms of a divisibility sequence is fundamental.

THEOREM III. *If p is a primitive divisor of u_n, and if μ is the restricted period of (u_n) modulo p, then $n | \mu$.*

Proof. Let $(n, \mu) = r$. Then there exist positive integers x and y such that $nx - \mu y = r$.

Since $u_n \equiv 0 \pmod{p}$, we have $u_{nx} \equiv 0 \pmod{p}$ (divisibility) $u_{nx} \equiv b^y u_{nx-\mu y} \pmod{p}$ (restricted period) or $u_{nx} \equiv b^y u_r \equiv 0 \pmod{p}$, whence $u_r \equiv 0 \pmod{p}$ as $b \not\equiv 0 \pmod{p}$ if $(a_k, p) = 1$. But as p is a primitive divisor of u_n, $u_r \equiv 0 \pmod{p}$ for $r \mid n$ implies $r = n$. Hence $(n, \mu) = r = n$, and $n \mid \mu$ as was to be shown.

Combining this with the information on μ given in § 2, it is seen that p is restricted to certain arithmetic progressions $tn + r_i$. For example, if the sequence is of second order $\mu \mid p - 1$ or $p + 1$, whence $p = tn \pm 1$.

4. *Divisibility Sequences of Third Order.* The condition $u_0 = 0$ makes it easy to find the regular divisibility sequences of first and second order. There is no regular sequence of first order unless the trivial sequence of zeros be considered a divisibility sequence. For second order we have $u_n = t(\alpha_1^n - \alpha_2^n)/(\alpha_1 - \alpha_2)$ or tna^{n-1} according as the roots of the associated polynomial are distinct or equal. The first of these is the well known Lucas sequence.

The consideration of third order sequences is by no means so simple. We may construct formally $u_n = \left(\dfrac{\alpha_1^n - \alpha_2^n}{\alpha_1 - \alpha_2}\right)^2 = \dfrac{\alpha_1^{2n} + \alpha_2^{2n} - 2\alpha_1^n\alpha_2^n}{\alpha_1^2 + \alpha_2^2 - 2\alpha_1\alpha_2}$ which satisfies a third order sequence whose characteristic polynomial has roots α_1^2, α_2^2, and $\alpha_1\alpha_2$. This will be a sequence of integers if $v_n = (\alpha_1^n - \alpha_2^n)/(\alpha_1 - \alpha_2)$ is of either the Lucas or Lehmer type. Such a sequence is essentially of quadratic type and there is nothing to be gained by considering it as a third order sequence. It is probable that there are no regular third order sequences of any other type.[7]

It is easily seen that we cannot obtain a divisibility sequence of third order satisfying an arbitrary recurrence merely by an appropriate choice of initial values. Consider $u_{n+3} = u_{n+1} + u_n$. From § 3 we must take $u_0 = 0$, and may take $u_1 = 1$. The condition $u_2 \mid u_4$ implies $u_2 = \pm 1$, but in neither case does $u_4 \mid u_8$.

If a sequence is of type v_n^2 as given above, its characteristic cubic $f(x)$ has a rational root $a = \alpha_1\alpha_2$. Hence if there is a third order divisibility sequence whose $f(x)$ is irreducible, it is certainly not of type v_n^2. This possibility is considered in the following theorem.

THEOREM IV. *There is no regular divisibility sequence (u_n), whose*

[7] Since completing this paper I have learned from Dr. Morgan Ward that he has been able to show that this is the only type if $f(x)$, the characteristic polynomial, has a linear and an irreducible quadratic factor. As this paper covers the case $f(x)$ irreducible, the only doubtful possibility is that $f(x)$ is the product of three linear factors.

characteristic polynomial is an irreducible cubic whose last two coefficients are relatively prime.

As the proof of this theorem is quite long, it will be subdivided into Lemmas. Lemma 4 gives the first of the equations which lead to the contradiction of the assumption that there is a divisibility sequence satisfying the requirements of the theorem.

Assume that there is a regular divisibility sequence (u_n), whose characteristic is $f(x) = x^3 - a_1x^2 - a_2x - a_3 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ where $(a_2, a_3) = 1$ and let $f(x)$ be irreducible. (u_n) satisfies the recurrence

$$u_{n+3} = a_1u_{n+2} + a_2u_{n+1} + a_3u_n. \tag{5}$$

As $f(x)$ is irreducible α_1, α_2, and α_3 are distinct and

$$u_n = c_1\alpha_1{}^n + c_2\alpha_2{}^n + c_3\alpha_3{}^n. \tag{6}$$

We note that as (u_n) is a regular divisibility sequence $u_0 = 0$ or

$$c_1 + c_2 + c_3 = 0. \tag{7}$$

Moreover we take $u_1 = 1$, as is permissable.

LEMMA 1. *If $p\,|\,a_3$, and $p\,|\,u_n$, then n has a factor r, $1 < r < \bar{n}$, $\bar{n}$ a fixed number.*

For if p divides any terms of (u_n), let u_m be the first. It evidently suffices to show $(m, n) \neq 1$. Then take $\bar{n}$ greater than m. As there are only a finite number of primes dividing a_3, there is one value for $\bar{n}$ which will do for all divisors of a_3. In fact, it can be shown that $a_3{}^3$ will suffice. Now if $(m, n) = 1$, there are positive integers x and y such that $mx = ny + 1$. By the divisibility property $u_{mx} \equiv 0 \pmod{p}$ and $u_{ny} \equiv 0 \pmod{p}$. From (5)

$$u_{mx} = a_1u_{mx-1} + a_2u_{mx-2} + a_3u_{mx-3}.$$

Now $p\,|\,u_{mx}$, $p\,|\,u_{mx-1} = u_{ny}$, $p\,|\,a_3$, but $p \nmid a_2$ as $(a_2, a_3) = 1$. Hence $p\,|\,u_{mx-2}$. Similarly as

$$u_{mx-1} = a_1u_{mx-2} + a_2u_{mx-3} + a_3u_{mx-4},$$

we have $p\,|\,u_{mx-3}$. Proceeding thus we finally obtain $p\,|\,u_1 = 1$, which is a contradiction. Hence $(m, n) \neq 1$.

LEMMA 2. *If $p\,|\,u_n$ and p is a divisor of the discriminant of $f(x)$, n has a factor less than a finite limit $\bar{n}$.*

If p also divides a_3 then Lemma 1 proves this. If $p \nmid a_3$, then p is either a primitive divisor of u_n or of u_r where $r|n$. In this case $r|\mu$ the restricted period of (u_n) modulo p, by reason of Theorem III, and $r \neq 1$ as $u_1 = 1$. As $f(x)$ is irreducible, its discriminant is not zero and has only a finite number of divisors. The restricted periods of these primes will lie below a finite limit $\bar{n}$. Hence $r < \bar{n}$ and so n has a factor less than $\bar{n}$.

LEMMA 3. *If q is a prime greater than $\bar{n}$, then* $u_q{}^6 \equiv u_1{}^6 \pmod{q}$, $u_{q^2}{}^6 \equiv u_1{}^6 \pmod{q}$.

By Lemma 1, u_q has no prime factor dividing a_3. As $u_1 = 1$, every prime dividing u_q is a primitive divisor of u_q. Hence if $p|u_q$ and μ is the restricted period of (u_n) modulo p, then $q|\mu$ by Theorem III. As p does not divide the discriminant of $f(x)$ by Lemma 2, we have $\mu | p-1$, p^2-1, or p^3-1, and hence $q|p^6-1$. Since $p^6 \equiv 1 \pmod{q}$ for every prime p dividing u_q, it follows by multiplication that $u_q{}^6 \equiv 1 \pmod{q}$ or $u_q{}^6 \equiv u_1{}^6 \pmod{q}$ as $u_1 = 1$. Now $p^6 \equiv 1 \pmod{q^2}$ for the primitive divisors of u_{q^2}, and hence à fortiori $p^6 \equiv 1 \pmod{q}$. Since all the divisors of u_{q^2} are primitive divisors of either u_q or u_{q^2}, we have $u_{q^2}{}^6 \equiv 1 \pmod{q}$ or $u_{q^2}{}^6 \equiv u_1{}^6 \pmod{q}$ as before.

LEMMA 4.

$$c_1\alpha_2 + c_2\alpha_3 + c_3\alpha_1 = \epsilon_1(c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3)$$

and

$$c_1\alpha_3 + c_2\alpha_1 + c_3\alpha_2 = \epsilon_2(c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3)$$

where $\epsilon_1{}^6 = \epsilon_2{}^6 = 1$.

For if q is a prime greater than $\bar{n}$, by Lemma 3 we have

$$(8) \qquad (c_1\alpha_1{}^q + c_2\alpha_2{}^q + c_3\alpha_3{}^q)^6 \equiv (c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3)^6 \pmod{q}.$$

Now if $f(x)$ is irreducible $\pmod{q}$ then

$$(9) \qquad \alpha_1{}^q \equiv \alpha_2, \ \alpha_2{}^q \equiv \alpha_3, \ \alpha_3{}^q \equiv \alpha_1 \pmod{Q}$$

where Q is a prime ideal dividing q in $K(\alpha_1, \alpha_2, \alpha_3)$. Hence from (8)

$$(10) \qquad (c_1\alpha_2 + c_2\alpha_3 + c_3\alpha_1)^6 \equiv (c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3)^6 \pmod{Q}.$$

Now if $f(x)$ is irreducible there are infinitely many primes q for which $f(x)$ is irreducible $\pmod{q}$.[8] Hence the difference of the two sides of (10) is an

[8] Hasse, "Bericht über Neuere Untersuchungen und Probleme aus der Theorie der Algebraischen Zahlkörper," Part II, p. 127, *Jahresbericht Ergänzungsbände,* vol. 6 (1930). Here K $(\alpha_1, \alpha_2, \alpha_3)$ is a cyclic extension of either the rational field or a quadratic field.

algebraic number divisible by infinitely many prime ideals, and consequently must be zero. Hence

$$c_1\alpha_2 + c_2\alpha_3 + c_3\alpha_1 = \epsilon_1(c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3) \tag{11}$$

where $\epsilon_1{}^6 = 1$. Similarly since $\alpha_1{}^{q^2} \equiv \alpha_3$, $\alpha_2{}^{q^2} \equiv \alpha_1$, $\alpha_3{}^{q^2} \equiv \alpha_2 \pmod{Q}$ and reasoning on u_{q^2} we have

$$c_1\alpha_3 + c_2\alpha_1 + c_3\alpha_2 = \epsilon_2(c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3) \tag{12}$$

where $\epsilon_2{}^6 = 1$.

Combining (7), (11) and (12) we have the system of equations:

$$\begin{array}{lllll} & c_1 & + c_2 & + c_3 & = 0 \\ (13) & c_1(\alpha_2 - \epsilon_1\alpha_1) & + c_2(\alpha_3 - \epsilon_1\alpha_2) & + c_3(\alpha_1 - \epsilon_1\alpha_3) & = 0 \\ & c_1(\alpha_3 - \epsilon_2\alpha_1) & + c_2(\alpha_1 - \epsilon_2\alpha_1) & + c_3(\alpha_2 - \epsilon_2\alpha_3) & = 0 \end{array}$$

If the c's all vanish, then the sequence (u_n) will consist merely of 0's. If not, the determinant of the c's

$$-(1 + \epsilon_1 + \epsilon_2)(\alpha_1{}^2 + \alpha_2{}^2 + \alpha_3{}^2 - \alpha_1\alpha_2 - \alpha_1\alpha_3 - \alpha_2\alpha_3)$$

must vanish.

If $\alpha_1{}^2 + \alpha_2{}^2 + \alpha_3{}^2 - \alpha_1\alpha_2 - \alpha_1\alpha_3 - \alpha_2\alpha_3 = 0$, then $a_1{}^2 = 3a_2$ and the roots of $f(x)$ are

$$\begin{array}{ll} & \alpha_1 = -a_1/3 + (a_1{}^3/27 - a_3)^{1/3} \\ (14) & \alpha_2 = -a_1/3 + \rho(a_1{}^3/27 - a_3)^{1/3} \\ & \alpha_3 = -a_1/3 + \rho^2(a_1{}^3/27 - a_3)^{1/3} \end{array}$$

where ρ is a primitive cube root of unity. Here for primes $q = 3k + 2$, $\alpha_1{}^q \equiv \alpha_1$, $\alpha_2{}^q \equiv \alpha_3$, $\alpha_3{}^q \equiv \alpha_2 \pmod{q}$ and reasoning as before

$$c_1 + c_2\rho^2 + c_3\rho = \epsilon_3(c_1 + c_2\rho + c_3\rho^2).$$

Trying the six possible values of ϵ_3, we find that two of the c's must be equal, or one must vanish. In no one of these cases can the sequence (u_n) be a sequence of rational integers.

If $1 + \epsilon_1 + \epsilon_2 = 0$, we have $\epsilon_1 = \rho$, $\epsilon_2 = \rho^2$. Solving (13) with

$$c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = u_1 = 1,$$

we obtain

$$c_1 = \frac{1}{\alpha_1 + \rho\alpha_2 + \rho^2\alpha_3},\quad c_2 = \frac{\rho}{\alpha_1 + \rho\alpha_2 + \rho^2\alpha_3},\quad c_3 = \frac{\rho^2}{\alpha_1 + \rho\alpha_2 + \rho^2\alpha_3}. \tag{15}$$

Here the vanishing of the denominators implies the vanishing of the second factor of the determinant, a possibility which has just been excluded. Here again the field is of the type $K(\sqrt[3]{d})$; for from the fact that u_2 is rational it is easily shown that $(\alpha_1 + \rho^2\alpha_2 + \rho\alpha_3)^3$ is rational. Hence for

$$q = 3k + 2,\ \alpha_1{}^q \equiv \alpha_1,\ \alpha_2{}^q \equiv \alpha_3,\ \alpha_3{}^q \equiv \alpha_2 \pmod{q}$$

and reasoning as before

$$c_1\alpha_1 + c_2\alpha_3 + c_3\alpha_2 = \epsilon_4(c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3).$$

Combining these six possibilities with (15) we have one of

$$\begin{array}{ll} \alpha_1 = \alpha_2 & 2\alpha_1 - \alpha_2 - \alpha_3 = 0 \\ \alpha_1 = \alpha_3 & 2\alpha_2 - \alpha_1 - \alpha_3 = 0 \\ \alpha_2 = \alpha_3 & 2\alpha_3 - \alpha_1 - \alpha_2 = 0 \end{array}$$

Each one of these contradicts the irreducibility of $f(x)$. For an irreducible polynomial has no equal roots, and if (say) $2\alpha_1 - \alpha_2 - \alpha_3 = 0$ then $3\alpha_1 = \alpha_1 + \alpha_2 + \alpha_3 = -a_1$, and the root α_1 is rational. This completes the proof of Theorem IV.

YALE UNIVERSITY.

THE CONSTRUCTION OF A NORMAL BASIS IN A SEPARABLE NORMAL EXTENSION FIELD.

By Ruth Stauffer.

1. *Introduction.* If K is a separable normal extension field over k and if the characteristic of k is not a divisor of the degree of K/k, then the group ring defined by the Galois Group, $\mathfrak{G}$, of K/k and the field k is a direct sum of simple algebras. These different simple algebras determine corresponding components for every element of K. The components of the elements of a basis of K/k can thus be arranged in matric form such that the components of one element form one column. From every column of this matrix it is possible to choose a non-vanishing component. The sum of these components, then, is an element of K which generates a *normal basis* of K/k, that is, it together with its conjugates forms a basis of K/k.

E. Noether (10), and A. Speiser (12) obtained results assuming the existence[1] of a normal basis. Noether formulated the problem in terms of operator isomorphisms. Speiser employed the theory of group representations. The formulae used by Speiser are obtained in this paper in some preliminary work on complementary bases. These same methods are then used to determine the discriminant of the centrum of the integral group ring. Since this discriminant is integral, a relation is obtained between the order of the group, the degree of the irreducible representations, and the class numbers.

2. *Complementary bases.* Two bases $a_i, \bar{a}_k$ of the semi-simple algebra $\mathfrak{o}$ (3), (5), (8), (13), (14) are said to be *complementary* if the trace matrix is the unit matrix, that is, $(Tr(a_i\bar{a}_k)) = E$. The following two theorems are direct consequences of this definition and of the linear property of the trace.

Theorem 1. *If $a_i\bar{a}_k$ are complementary bases and $b_i, \bar{b}_k$ are also bases of $\mathfrak{o}$, then there exist non-singular matrices P, Q in Ω the fundamental field of $\mathfrak{o}$, such that*

$$(\cdot\ \cdot \bar{a}_k \cdot\ \cdot) = (\cdot\ \cdot \bar{b}_k \cdot\ \cdot)P, \qquad \begin{bmatrix} \cdot \\ a_i \\ \cdot \end{bmatrix} = Q \begin{bmatrix} \cdot \\ b_i \\ \cdot \end{bmatrix},$$

and b_i and $\bar{b}_k$ are complementary if and only if $QP = E$.

[1] The existence of a normal basis for Galois fields was proved by Hensel in 1888, (7). Deuring (4), Hasse (6), and Brauer (1) have proved the existence of a normal

THEOREM 2. *If both $a_i, \bar{a}_k$ and $b_i, \bar{b}_k$ are complementary bases of $\mathfrak{o}$, and Γ, Λ are any two representations of $\mathfrak{o}$, then*

$$\sum_i Tr_\Gamma(a_i) Tr_\Lambda(\bar{a}_i) = \sum_i Tr_\Gamma(b_i) Tr_\Lambda(\bar{b}_i), \tag{2.1}$$

where $Tr_\Gamma(a_i)$ means the trace with respect to the representation Γ of a_i.

We note from Theorem 1 that every basis has a complementary basis. The following are outstanding examples of such bases: 1) the group elements $(\cdot\cdot S_i \cdot\cdot)$ and $(\cdot\cdot S_i^{-1}/g \cdot\cdot)$ of the group ring (13) where g is the order of the group, 2) the matric units, (8), (13), $(\cdot\cdot c_{ik} \cdot\cdot)$ and $(\cdot\cdot c_{ki}/n \cdot\cdot)$ of a total matric ring of degree n, 3) the matric units $(\cdot\cdot c_{ik} \cdot\cdot)$ and $(\cdot\cdot c_{ki}^{(l)}/f_l \cdot\cdot)$ of a semi-simple ring in which the coefficient field is algebraically closed, and f_l is the degree of the simple algebra defined by $\{c_{ik}^{(l)}\}$.

3. *Orthogonality relations.* We shall denote an irreducible representation of the semi-simple algebra $\mathfrak{o}$, by Γ_μ, and the trace with respect to Γ_μ, by χ_μ. If, then, a_i and $\bar{a}_k$ are complementary bases of $\mathfrak{o}$, the equality (2.1) may be written as

$$\sum_i \chi_\nu(a_i)\chi_\mu(\bar{a}_k) = \sum_{i,k,l} \chi_\nu(c_{ik}^{(l)})\chi_\mu(c_{ki}^{(l)}/f_l).$$

Since $\chi_\nu(c_{ik}^{(l)}) = 0$ for $\nu \neq l$ and $= \delta_{ik}$ for $\nu = l$, we conclude that

$$\sum_i \chi_\nu(a_i)\chi_\mu(\bar{a}_i) = \delta_{\nu\mu}, \tag{3.1}$$

our *first set of orthogonality relations.*

We note that if $\mathfrak{o}$ is a group ring and the elements of the group are S_i, then the complementary basis is S_i^{-1}/g and

$$\sum_i \chi_\mu(S_i)\chi_\nu(S_i^{-1}) = g\delta_{\mu\nu}, \tag{3.2}$$

if the characteristic of the fundamental field does not divide g.

In the case of a group ring there is also an orthogonality property for the sums of the different irreducible representations of a fixed element of the group and its inverse. Consider the *classes of conjugates of a group.* We shall denote by K_S the sum of the elements of a class generated by S. The elements K_S are commutative with all the group elements, and, moreover, generate the centrum, (8), p. 692. Furthermore, if $ASA^{-1} = T$, then $AS^{-1}A^{-1} = T^{-1}$ and if $S = T$, then $S^{-1} = T^{-1}$. Therefore the number of elements in K_S is

---

basis of K/k when k has a zero characteristic. In Deuring's paper the proof has been extended to the general case that k have any characteristic.

Numbers in parentheses refer to the bibliography on page 597.

equal to the number of elements in $K_{S^{-1}}$. If, then, h_i denotes the number of elements of the class of S_i,

$$\sum_i h_i\chi_\mu(T_i)h_i\chi_\nu(T_i^{-1}/h_ig) = \sum_i \chi_\mu(S_i)\chi_\nu(S_i^{-1}/g) = \delta_{\mu\nu},$$

where T_i is a representative of K_{S_i}. Therefore

$$\sum_i \chi_\mu(K_{S_i})\chi_\nu(K_{S_i^{-1}}/h_ig) = \delta_{\mu\nu}, \tag{3.3}$$

or

$$(\chi_\mu(K_{S_i}))(\chi_\nu K_{S_k^{-1}}/h_kg)) = E,$$

where μ, k denote rows, i, ν denote columns. Therefore

$$(\chi_\nu(K_{S_k^{-1}}/h_kg))(\chi_\mu(K_{S_i})) = E.$$

That is,

$$\sum_\mu \chi_\mu(K_{S_k^{-1}}/h_kg)\,\chi_\mu(K_{S_i}) = \delta_{ik}, \tag{3.4}$$

our *second set of orthogonality relations.*

4. *Discriminant of the centrum of the group ring with integral coefficients.* If the irreducible representation Γ_ν is of degree f_ν, then

$$\chi_\nu(K_{S_k^{-1}}/h_kg)\chi_\nu(K_{S_i}) = f_\nu\chi_\nu(K_{S_k^{-1}}/h_kg\cdot K_{S_i})$$

and therefore

$$\sum_\nu f_\nu\chi_\nu(K_{S_k^{-1}}/h_kg\cdot K_{S_i}) = \delta_{ik}.$$

Hence, if Γ is the regular representation of the group ring,

$$Tr_\Gamma(K_{S_i}K_{S_k^{-1}}/h_kg) = \sum_\nu f_\nu\chi_\nu(K_{S_i}K_{S_k^{-1}}/h_kg) = \delta_{ik}. \tag{4.1}$$

That is, the elements $K_{S_i^{-1}}/h_ig$ form a basis of the centrum which is "complementary" to the basis K_{S_i}, "complementary" in the sense that the traces are traces with respect to the regular representation of the group.

Let the group ring $\mathfrak{G}(\Omega)$ be defined by the group $\mathfrak{G}$ and the field Ω. Then to determine the *discriminant, d, of the centrum,* write $\mathfrak{G}(\Omega)$ as the direct sum of two-sided simple ideals $\mathfrak{a}^{(i)}$. Let e_i be the corresponding components of the unit element, e, of $\mathfrak{G}(\Omega)$. Then $e_1, \cdots, e_r$ form a basis for the centrum, and the discriminant with respect to this basis is $|Tr(e_ie_k)| = 1$. The trace, of course, is taken with respect to the regular representation of the centrum. The determinant of the traces with respect to the regular representation of the group ring is

$$(4.2) \qquad | Tr_\Gamma(e_ie_k) | = \prod_{i=1}^{r} f_i^2.$$

If $\begin{bmatrix} \cdot \\ K_{S_i} \\ \cdot \end{bmatrix} = P \begin{bmatrix} \cdot \\ e_i \\ \cdot \end{bmatrix}$, and $(\cdot \cdot K_{S_i^{-1}}/h_i g \cdot \cdot) = (\cdot \cdot e_i \cdot \cdot)Q$, then

$$| Tr_\Gamma(K_{S_i}K_{S_k^{-1}}/h_k g) | = P \,| Tr_\Gamma(e_ie_k) |\, Q,$$

and

$$| Tr(K_{S_i}K_{S_k^{-1}}/h_k g) | = P \,| Tr(e_ie_k) |\, Q.$$

We conclude, from (4.1) and (4.2) that

$$\frac{d}{g^r \Pi h_i} = \frac{| Tr(K_{S_i}K_{S_k^{-1}}/h_k g) |}{| Tr(e_ie_k) |} = \frac{| Tr_\Gamma(K_{S_i}K_{S_k^{-1}}/h_k g) |}{| Tr_\Gamma(e_ie_k) |} = \frac{1}{\Pi f_i^2}.$$

Moreover, d is an integer, for the elements K_{S_i} are group ring elements with integral coefficients. It follows that $g^r \Pi h_i$ is divisible by Πf_i^2.

5. *Orthogonality relations of the coefficients of the representations.* If $\mathfrak{o}$ is a semi-simple algebra, a_i, $\bar{a}_k$ complementary bases of $\mathfrak{o}$, b_i, $\bar{b}_k$ also complementary bases of $\mathfrak{o}$, and if Γ, Λ are any two representations of $\mathfrak{o}$, then it follows from the linear property of traces and from the definition of complementary bases that

$$(5.1) \qquad \sum \alpha^{(i)}_{lm,\gamma} \alpha^{(i)}_{sr,\lambda} = \sum \beta^{(i)}_{lm,\gamma} \beta^{(i)}_{sr,\lambda}$$

where

$$\Gamma(a_i) = (\alpha^{(i)}_{lm,\gamma}),\ \Lambda(a_i) = (\alpha^{(i)}_{lm,\lambda}),\ \text{and}\ \Gamma(b_i) = (\beta^{(i)}_{lm,\gamma}),\ \Lambda(b_i) = (\beta^{(i)}_{lm,\lambda}).$$

On calculating the sum (5.1) for the representation of the complementary bases $(c^{(l)}_{ik})$, $(c^{(l)}_{ki}/f_l)$ of $\mathfrak{o}$, we see that

$$\sum \alpha^{(i)}_{lm,\gamma} \alpha^{(i)}_{sr,\lambda} = 0, \qquad \lambda \neq \gamma$$

and for $\lambda = \gamma$,

$$(5.2) \qquad \sum \alpha^{(i)}_{lm,\gamma} \alpha^{(i)}_{sr,\lambda} = \begin{cases} 0, & (l,m) \neq (r,s) \\ 1/f, & (l,m) = (r,s) \end{cases}$$

We note that the orthogonality property of the coefficients of the representations of the group elements for the group ring is a particular case of relations (5.2).

6. *Calculation of the matric units by means of complementary bases.* Another problem to which complementary bases may be applied is that of calculating the matric units of the group ring $\mathfrak{G}(\Omega)$, where Ω is an algebraically closed field for which the characteristic is not a divisor of the order

of the group. We write $\mathfrak{G}(\Omega) = \mathfrak{a}^{(1)} + \cdots + \mathfrak{a}^{(r)}$, where $\mathfrak{a}^{(l)}$ are two-sided simple ideals. Let the r distinct irreducible representations of $\mathfrak{G}$ be

$$s \rightarrow S^{(l)} = (\sigma_{ik}^{(l)}), \qquad t \rightarrow T^{(l)} = (\tau_{ik}^{(l)}), \cdots, \qquad (l = 1, 2, \cdots, r).$$

It follows from Theorem 1 that

$$(e, s, t, \cdots) = (\cdots c_{ik}^{(l)} \cdots)P, \text{ and } \begin{bmatrix} \cdot \\ c_{ki}^{(l)}/f_l \\ \cdot \end{bmatrix} = P \begin{bmatrix} \cdot \\ s^{-1}/g \\ \cdot \end{bmatrix},$$

where g is the order of the group, and f_l the degree of the representation defined by $\mathfrak{a}^{(l)}$. Therefore, in order to determine the matric units we need only determine P. From the theory of representations we know

$$s = \sum_{ik} c_{ik}^{(1)} \sigma_{ik}^{(1)} + \cdots + \sum_{ik} c_{ik}^{(r)} \sigma_{ik}^{(r)}.$$

Hence

$$P = \begin{bmatrix} \cdot & \cdot & \cdot \\ \epsilon_{ik}^{(l)}, & \sigma_{ik}^{(l)}, & \tau_{ik}^{(l)}, \\ \cdot & \cdot & \cdot \end{bmatrix}$$

and

$$P \begin{bmatrix} \cdot \\ s^{-1} \\ \cdot \end{bmatrix} = g \begin{bmatrix} \cdot \\ c_{ki}^{(l)}/f_l \\ \cdot \end{bmatrix} = \begin{bmatrix} \epsilon_{ik}^{(l)} e^{-1} + \sigma_{ik}^{(l)} s^{-1} + \tau_{ik}^{(l)} t^{-1} + \cdots \end{bmatrix}.$$

In other words, if

(6.1) $$M^{(l)} = E^{(l)} e^{-1} + S^{(l)} s^{-1} + T^{(l)} t^{-1} + \cdots = (\mu_{ik}^{(l)}),$$

then

(6.2) $$g c_{ki}^{(l)}/f_l = \mu_{ik}^{(l)}.$$

This discussion can be applied to an arbitrary semi-simple algebra, if instead of the basis $(e, s, \cdots)$ and its complementary basis, one considers any set of complementary bases a_i and $\bar{a}_k$. In the general case the matrix is

(6.3) $$M^{(l)} = \sum_i \bar{A}_i^{(l)} a_i.$$

7. *Complementary bases for commutative normal fields. Idempotents.* If K and K are isomorphic separable normal extension fields of degree n over k, the direct product ring K_{K} consisting of the set of all elements $\epsilon = \Sigma d_i \delta_i$ where d_i belongs to K and δ_i belongs to K, is directly decomposible [(8), p. 683] into a sum of n simple ideals, $e_i K$,[2] where e_i are indecomposible components of unity. That is

(7.1) $$e_i e_k = \delta_{ik} e_i \quad \text{and} \quad \sum_i e_i = e,$$

[2] $e_i K$ is a field with unit element e_i and isomorphic to K.

where e is the unit element of K_{K}. Let $\alpha = \Sigma e_i a_i$, where $\alpha \epsilon \mathrm{K}$ and $a_i \epsilon K$. Then

$$e_i \alpha = e_i a_i. \tag{7.2}$$

That is e_i defines a representation of the first degree which maps α on a_i. Furthermore if e is any indecomposible component of unity, say e_1, then the set $\{e_i\}$ is actually the set of conjugates $\{e^S\}$ (4), p. 141.

We now choose $\bar{S}$, an element of the Galois group of K over k such that

$$e\alpha^{\bar{S}} = ea^S. \tag{7.3}$$

That is $\bar{S}$ is the automorphism of K which takes α into the element corresponding to a^S. It follows from (7.2) that $e^S\alpha = e^S a^S$, and from (7.3) that $e^{\bar{S}^{-1}}\alpha = e^{\bar{S}^{-1}}a^S$. In other words, e^S maps α on a^S and $e^{\bar{S}^{-1}}$ maps α on a^S. However, a was an arbitrary element of K. We may conclude, therefore, that $e^S = e^{\bar{S}^{-1}}$. We note that if $E, S, T, \cdots$ are the elements of the Galois Group, $\mathfrak{G}$, of K/k and $\bar{E}, \bar{S}, \bar{T}, \cdots$ are the elements of the Galois Group $\bar{\mathfrak{G}}$ of K/k, then $e^E, e^S, e^T, \cdots$ are distinct and hence $e^{\bar{E}}, e^{\bar{S}}, e^{\bar{T}}, \cdots$ are also distinct. Therefore the correspondence defined between $\mathfrak{G}$ and $\bar{\mathfrak{G}}$ is a one to one correspondence.

If z_i is a basis of K/k, hence also a basis of $K_{\mathrm{K}}/\mathrm{K}$ and if the idempotent $e = \Sigma z_i \alpha_i$, $e^{\bar{S}} = \Sigma z_i \alpha_i^{\bar{S}}$, then $\alpha_i^{\bar{S}}$ is that representation of the complementary basis of z_i defined by $e^{\bar{S}}$, (9), p. 538. This theorem is a key to the first step in the proofs of the existence of a normal basis, namely the proof that $\mathfrak{G}(\mathrm{K})$ is operator isomorphic to $K(\mathrm{K})$, where $\mathfrak{G}$, K, and K are as previously defined, and $K(\mathrm{K})$ is the operator module consisting of the elements of the ring K_{K}, the operators being the elements of the Galois Group $\mathfrak{G}$ of K applied as automorphisms of K. The proof is as follows. Given any basis $z_1, \cdots, z_n$ of K/k, we form the product ring K_{K} of K with an isomorphic extension field K of k. Then the matrix,

$$\begin{pmatrix} \alpha_1^{\bar{E}^{-1}}, & \alpha_1^{\bar{S}^{-1}}, & \cdots \\ \cdot & \cdot & \cdot\ \cdot \\ \cdot & \cdot & \cdot\ \cdot \\ \alpha_n^{\bar{E}^{-1}}, & \alpha_n^{\bar{S}^{-1}}, & \cdots \end{pmatrix}$$

where α_i are as defined above, transforms $z_1, \cdots, z_n$ into a set of linearly independent conjugates $e^E, \cdots, e^S, \cdots$, the indecomposible components of the unit element of K_{K}. The correspondence $S \leftrightarrow e^S$ then, defines the operator isomorphism

$$\mathfrak{G}(\mathrm{K}) \simeq K(\mathrm{K}).$$

This is essentially the proof given both by Deuring and by Hasse. Deuring works with the idempotents without stating the explicit formulae, Hasse works with the set of coefficients of the idempotents. These coefficients are, in reality, the columns of the matrix of transformation $(\alpha_i^{\bar{S}^{-1}})$.

8. *Construction of a normal basis.* We first construct a set of linearly independent elements which are isomorphic to the matric units of the group ring $\mathfrak{G}(\Omega)$, where Ω is an algebraically closed field over k, and $\mathfrak{G}$ is the Galois group of K/k. By so choosing Ω every irreducible representation of $\mathfrak{G}$ may be expressed by means of matrices with coefficients in Ω. Let $E, A, B, \cdots, S, \cdots$ be the elements of $\mathfrak{G}$, $z^E, z^A, \cdots, z^S \cdots$, the conjugates of z, and Γ an absolute irreducible representation of $\mathfrak{G}$ defined as follows:

$$E = (\epsilon_{ik}), \quad A = (\alpha_{ik}), \cdots \; S = (\sigma_{ik}), \cdots. \tag{8.1}$$

Let

$$M(z) = Ez^{E^{-1}} + Az^{A^{-1}} + \cdots + Sz^{S^{-1}} + \cdots = (\sum_{\mathfrak{G}} \sigma_{ik} z^{\;-1}), \tag{8.2}$$

where z is any element of K. If z generates a normal basis of K/k, the matrix $M(z)$ corresponds to the matrix (6.1) in the isomorphism between $\mathfrak{G}(k)$ and $K(k)$. For convenience let

$$\sum_{\mathfrak{G}} \sigma_{ik} z^{S^{-1}} = \xi_{ik}(z). \tag{8.3}$$

Thus $M(z) = (\xi_{ik}(z))$. We first show that there exists an element z of K such that the $\xi_{ik}(z)$ are linearly independent. For this purpose we consider representations of $\mathfrak{G}$ defined by the $\xi_{ik}(z)$. If $S^{-1} = (\bar{\sigma}_{ik})$,[3]

$$(\xi_{i1}(z^{S^{-1}}), \cdots, \xi_{if}(z^{S^{-1}})) = (\xi_{i1}(z), \cdots, \xi_{if}(z))(\sigma_{ik}), \tag{8.4}$$

This equality is true for $i = 1, 2, \cdots, f$, where f is the degree of Γ. Thus $(\bar{\sigma}_{ik})$ is a representation of S^{-1} on Ω defined by the representation module $(\{\xi_{i1}(z)\}, \cdots, \{\xi_{if}(z)\})$, where the sets of columns $\{\xi_{ik}(z)\}$, $i = 1, \cdots, f$ are considered as elements of the module. Similarly,

$$(\xi_{1i}(z)^{S^{-1}}, \cdots, \xi_{fi}(z)^{S^{-1}}) = (\xi_{1i}(z), \cdots, \xi_{fi}(z))(\bar{\sigma}_{ik}). \tag{8.4'}$$

Hence $(\bar{\sigma}_{ik})$ is a representation of S^{-1} on Ω defined by the representation module $(\{\xi_{1i}(z)\}, \cdots, \{\xi_{fi}(z)\})$. However, $(\bar{\sigma}_{ik})$ is an irreducible representation of S^{-1} on Ω defined by Γ. Therefore the representation modules $(\{\xi_{i1}(z)\}, \cdots, \{\xi_{if}(z)\})$ and $(\{\xi_{1i}(z)\}, \cdots, \{\xi_{fi}(z)\})$ must be zero or simple, in fact, in the latter case, $\{\xi_{i1}(z)\}, \cdots, \{\xi_{if}(z)\}$ and also $\{\xi_{1i}(z)\}, \cdots, \{\xi_{fi}(z)\}$ must be linearly independent. Furthermore,

---

[3] $(\bar{\sigma}_{k_i})$ is the adjoint representation, $\bar{\Gamma}(S)$, see (11), p. 163.

LEMMA 1. *There exists an element z of K such that $\xi_{ik}(z) \neq 0$ for every (i, k).*

For suppose that $\xi_{ik}(z) = 0$ for every z belonging to K, and for (i, k) fixed, then if $z_1, \cdot \cdot, z_n$ is a basis of K/k, the discriminant of K with respect to this basis is

$$D_z = \begin{vmatrix} z_1 & \cdots & z_n \\ \cdots & \cdots & \cdots \\ z_1^S & \cdots & z_n^S \\ \cdots & \cdots & \cdots \end{vmatrix}^2 = 0,$$

but $D_z \neq 0$ since K is separable over k. We may conclude, therefore, that there exists a z of the z_i such that $\xi_{ik}(z) \neq 0$, (i, k) fixed. Then $(\xi_{i1}(z), \cdot \cdot, \xi_{if}(z)) \neq 0$, and therefore $\xi_{i1}(z), \cdot \cdot \xi_{if}(z)$ are linearly independent. This demands that they all be different from zero, hence that all the modules $(\xi_{1i}(z), \cdot \cdot, \xi_{fi}(z))$ be different from zero. Thus all $\xi_{ik}(z) \neq 0$. Also,

LEMMA 2. *If $\mathfrak{G}(\Omega) = \mathfrak{a}^{(1)} + \cdot \cdot + \mathfrak{a}^{(r)}$, $\mathfrak{a}^{(i)}$ two-sided simple ideals, then the basis of every simple right ideal of $\mathfrak{a}$*[4] *can be written as a linear combination of the bases of simple right ideals $\mathfrak{r}_i$ defined by the matric units c_{ik},*

$$\{\mathfrak{r}_i\} = (c_{i1}, \cdot \cdot \cdot, c_{if}).$$

That is there exist α in Ω such that

$$\{\mathfrak{r}\} = \alpha\{c_{1i}\} + \cdot \cdot \cdot + \alpha_f\{c_{fi}\}. \tag{8.5}$$

For every simple right ideal of $\mathfrak{a}$ is isomorphic to $\mathfrak{r}_1$, and, hence, there exists an element a in the left ideal $(c_{11}, \cdot \cdot, c_{f1})$ such that

$$\{\mathfrak{r}\} = a\{\mathfrak{r}_1\} = (\sum_i \alpha_{i1} c_{i1}) \{\mathfrak{r}_1\} = \sum \alpha_{i1}\{\mathfrak{r}_i\}.$$

THEOREM 3. *The elements $\xi_{ik}(z)$ defined by* (8.3) *and which satisfy Lemma* 1, *are linearly independent.*

Let $\mathfrak{G}(\Omega) = \mathfrak{a}^{(1)} + \cdot \cdot + \mathfrak{a}^{(r)}$, and let $\mathfrak{r}$, a simple right ideal in $\mathfrak{a} = \mathfrak{r}_1 + \cdot \cdot + \mathfrak{r}_f$, define the irreducible representation Γ, (8.1). Thus $\mathfrak{a}$, is a representation module defining the representation

[4] $\mathfrak{a}$ may be any of the two sided-simple ideals $\mathfrak{a}^{(i)}$, and c_{ik} the matric units of $\mathfrak{a}$.

$$\begin{bmatrix} \cdot & & & 0 \\ & \cdot & & \\ & & (\sigma_{ik}) & \\ & & & \cdot \\ 0 & & & \cdot \end{bmatrix} \text{ of } S, \text{ and } \begin{bmatrix} \cdot & & & 0 \\ & \cdot & & \\ & & (\bar{\sigma}_{ik}) & \\ & & & \cdot \\ 0 & & & \cdot \end{bmatrix} \text{ of } S^{-1}.$$

On the other hand, M is a representation module defining the same representation of S^{-1}. Therefore $\mathfrak{a}$ is operator homomorphic to M, i. e., $\mathfrak{a} \to M$, and

$$c_{ik} \to \xi_{ki},$$
$$(8.6) \qquad c_{ik}\Omega \to \xi_{ki}\Omega,$$
$$c_{ik}S \to \xi^S_{ki}.$$

Furthermore, M is operator isomorphic to a factor group $\mathfrak{a}/\mathfrak{r}$ where $\mathfrak{r}$ is a right ideal in $\mathfrak{a}$ and corresponds to the zero element in M. Since the ring is completely reducible $\mathfrak{r} = \Sigma \mathfrak{r}_i^*$, where $\mathfrak{r}_i^*$ are simple right ideals. It follows from Lemma 2, (8.5), that $\mathfrak{r}_1^* = \Sigma \alpha_{i1}\{\mathfrak{r}_i\}$, and if we apply the homomorphism (8.6) to this equality, we may conclude that

$$0 = \Sigma \alpha_{k1}\{\xi_{ik}\} = \alpha_{11}\{\xi_{i1}\} + \cdot\cdot\cdot + \alpha_{f1}\{\xi_{if}\}.$$

The sets $\{\xi_{i1}\}, \cdot\cdot, \{\xi_{if}\}$, however, define a non-zero simple module and are linearly independent with respect to Ω. Therefore $\alpha_{i1} \equiv 0$. Similarly it can be shown that the coefficients of $\mathfrak{r}_i^*$ all vanish. Hence

$$\mathfrak{r} = \mathfrak{r}_1^* + \cdot\cdot\cdot + \mathfrak{r}_s^* = 0,$$

and the homomorphism is an isomorphism. We may conclude, therefore, that $\xi_{ik}(z)$ are linearly independent with respect to Ω.

For each of the r distinct irreducible representations of $\mathfrak{G}$ defined by the two-sided simple ideals $\mathfrak{a}^{(l)}$ of $\mathfrak{G}(\Omega)$, we now define a matrix

$$M^{(l)} = (\xi^{(l)}_{ik}(z_l)),$$

where $\xi^{(l)}_{ik}(z_l) \neq 0$ for all i, k. It remains to determine an element w belonging to K such that $\xi^{(l)}_{ik}(w) \neq 0$ for all i, k, l. Let the centrum idempotents of $\mathfrak{G}(\Omega)$ in $\mathfrak{a}^{(1)}, \cdot\cdot, \mathfrak{a}^{(r)}$ be $E^{(1)}, \cdot\cdot, E^{(r)}$ respectively, and

$$z^{TE^{(i)}} = \sum_{\mathfrak{G}} \alpha_s z^{TS}, \quad \text{if} \quad E^{(i)} = \sum_{\mathfrak{G}} \alpha_S S.$$

We construct

$$(8.7) \qquad w = z_1^{E^{(1)}} + \cdot\cdot\cdot + z_r^{E^{(r)}},$$

where z_l have been chosen such that $\xi_{ik}^{(l)}(z_l) \neq 0$. Furthermore, if $\mathfrak{a}^{(l)}$ is conjugate to $\mathfrak{a}^{(m)}$ with respect to k, we take $z_l = z_m$. This is possible since $\xi_{ik}^{(m)}(z_l)$ is conjugate to $\xi_{ik}^{(l)}(z_l)$ and therefore is not equal to zero if $\xi_{ik}^{(l)}(z_l) \neq 0$. If k contains the n-th roots of unity the case that $\mathfrak{a}^{(l)}$ and $\mathfrak{a}^{(m)}$ are conjugate $(l \neq m)$ does not occur.

Does w, defined as the sum of these components, belong to K? If k contains the n-th roots of unity, $z^{E^{(l)}}$ and hence w belong to K. If k does not contain the n-th roots of unity, we consider the sums of the conjugate idempotents. These sums give us elements of $\mathfrak{G}(k)$. That is,

$$e^{(i)} = \sum_{\text{Conjugates}} E^{(i)} = \sum_{\mathfrak{G}} \alpha_{si} S, \qquad \alpha_{si} \in k. \tag{8.8}$$

Furthermore, since we have the same z_i for conjugate $E^{(i)}$, we can write w as the sum $\sum_i z^{e^{(i)}}$. Hence w belongs to K and is defined as a *rational sum.*

It must now be shown that $\xi_{ik}^{(l)}(w) \neq 0$ for every i, k, l. To do this we apply the homomorphisms of (8.6). It follows from the construction of w, (8.7), that

$$\begin{aligned} \xi_{ik}^{(l)}(w) &= \sum_{\mathfrak{G}} \sigma_{ik}^{(l)} w^S = \sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (\sum_m z_m^{E^{(m)}})^S \\ &= \sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (z^{E^{(1)}})^S + \cdots + \sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (z_l^{E^{(l)}}) + \cdots + \sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (z_r^{E^{(r)}})^S. \end{aligned} \tag{8.9}$$

The general term of the right hand side of (8.9) may be expressed

$$\sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (z_\mu^{E^{(\mu)}})^S = {}^{E^{(\mu)}}\xi_{ik}^{(l)}(z_\mu),$$

and corresponds, in the homomorphism defined by

$$c_{ki}^{(l)} \to \xi_{ik}^{(l)}(z_\mu), \tag{8.10}$$

to $E^{(\mu)} c_{ki}^{(l)} = 0$ if $\mu \neq l$. Thus

$$\sum_{\mathfrak{G}} \sigma_{ik}^{(l)} (z_\mu^{E^{(\mu)}}) = 0, \qquad \mu \neq l.$$

If, however, $\mu = l$, $E^l c_{ki}^{(l)} = c_{ki}^{(l)}$. Therefore, in this case,

$${}^{E^{(l)}}\xi_{ik}^{(l)}(z_l) = \xi_{ik}^{(l)}(z_l).$$

We conclude that

$$\xi_{ik}^{(l)}(w) = \xi_{ik}^{(l)}(z_l) \neq 0, \tag{8.11}$$

for any i, k, l. It follows, as shown previously, that $\xi_{ik}^{(l)}(w)$ are linearly independent with respect to Ω. Moreover, $\xi_{ik}^{(l)}(w)$ are expressed linearly by means

of the set of elements $\{w^S\}$. Therefore this set $\{w^S\}$ forms a *normal basis* of K/k.

9. *A rational construction.* In 8 w was defined as a *rational sum,* but the z_l were chosen to satisfy the *irrational condition* $\xi^{(l)}_{ik}(z_l{}^{E^{(l)}}) \neq 0$. This irrational condition, however, may be replaced by a rational condition, $z^{e^{(l)}} \neq 0$, where $e^{(l)}$ is as defined in equation (8.8). Since the regular representation $\Gamma = \Sigma f_l \Gamma_l$,

$$\sum_{i,l} (f_l/n)\sigma^{(l)}_{ii} = Tr(S/n) = \delta_{SE},$$

and therefore

$$\sum_{i,l} (f_l/n)\xi^{(l)}_{ii}(z) = \sum_{\mathfrak{G}} (\sum_{i,l} (f_l/n)\sigma^{(l)}_{ii}) z^S = z.$$

This relation, together with the fact (8.11) that

$$\xi^{(l)}_{ik}(z^{e^{(l)}}) = \xi^{(l)}_{ik}(z^{E^{(l)}}),$$

gives us the equivalence of the two conditions

$$\xi^{(l)}_{ik}(z^{E^{(l)}}) \neq 0, \quad \text{and} \quad z^{e^{(l)}} \neq 0.$$

First suppose

$$\xi^{(l)}_{ik}(z^{E^{(l)}}) = \xi^{(l)}_{ik}(z^{e^{(l)}}) \neq 0.$$

Hence in the isomorphism defined by

$$c^{(l)}_{ii} \leftrightarrow \xi^{(l)}_{ii}(z^{e^{(l)}}),$$

$$z^{e^{(l)}} = \sum_{i,l} (f_l/n)\xi^{(l)}_{ii}(z^{e^{(l)}}) \leftrightarrow \sum_{i,l} (f_l/n)c^{(l)}_{ii} = f_l/n \neq 0.$$

Therefore

$$z^{e^{(l)}} \neq 0.$$

Conversely, if $z^{e^{(l)}} \neq 0$, one of the summands of

$$\sum_{i,l} (f_l/n)\xi^{(l)}_{ii}(z^{e^{(l)}})$$

must be different from zero. We have shown, however, that if one $\xi^{(l)}_{ik}(z^{e^{(l)}}) \neq 0$, then $\xi^{(l)}_{ik}(z^{e^{(l)}}) \neq 0$, l fixed. Hence the two conditions are equivalent and the following theorem is proved.

THEOREM 4. *If z_i is any element of K such that its component $z^{e^{(l)}} \neq 0$, then $w = \Sigma z_i{}^{e^{(l)}}$ will generate a normal basis of K/k.*

10. *The most general normal basis.* The final problem of this paper is to determine from a given normal basis the most general normal basis. We

note first that if $\mathfrak{G}(k)$ is operator isomorphic to $K(k)$ with operator region $\mathfrak{G}$ and if $\alpha \leftrightarrow a$, $\beta \leftrightarrow b$, where $\alpha, \beta \in G(k)$ and $a, b \in K(k)$, then $a^{(\beta)} = {}^{(\alpha)}b$, where $a^{(\beta)}$ means that a is operated on by β from the right and ${}^{(\alpha)}b$ means that b is operated on by α from the left.

THEOREM 5. *If $(w^{S_1}, \cdot \cdot, w^{S_n})$ is any normal basis of K/k, then the most general normal basis of K/k is of the form $(w^{\nu S_1}, \cdot \cdot, w^{\nu S_n})$ where ν is any element of $\mathfrak{G}(k)$ which is not a zero divisor in $\mathfrak{G}(k)$.*

Let $E \leftrightarrow w$, then $S \leftrightarrow w^S$. Suppose that $v^{S_1}, \cdot \cdot, v^{S_n}$ is any other normal basis of K/k. In the isomorphism $\mathfrak{G}(k) \simeq K(k)$ let $\nu \leftrightarrow v$. Then we know from the above note that

$$(\nu S_1, \cdot \cdot, \nu S_n) \leftrightarrow (v^{S_1}, \cdot \cdot v^{S_n}) = (w^{\nu S_1}, \cdot \cdot, w^{\nu S_n}).$$

On the other hand, if ν is any element of $\mathfrak{G}(k)$, then

$$\nu(S_1, \cdot \cdot, S_n) \leftrightarrow (w^{\nu S_1}, \cdot \cdot, w^{\nu S_n}).$$

Furthermore suppose there exist $c_i \neq 0$ such that

$$\Sigma \nu c_i S_i \leftrightarrow \Sigma c_i w^{\nu S_i} = 0,$$

then if ν is not a divisor of zero, $\nu = 0$ since $\Sigma c_i S_i \neq 0$. Hence if ν is not a divisor of zero the set $(w^{\nu S_1}, \cdot \cdot, w^{\nu S_n})$ is a normal basis of K/k.

Note. R. Brauer has told me of another proof of Theorem 5. It is the following: Let w^G be a normal basis, then

$$(w^G)^T = \Sigma \delta_{GT,H} w^H. \tag{1*}$$

The matrix $S_T = (\delta_{GT,H})$, G gives the row, H the column, is the matrix corresponding to T in the regular representation. (1*) may then be written

$$w^T = S_T(w). \tag{2*}$$

Let z^G be a second normal basis and let

$$w = A(z),$$

that is, $w^G = \Sigma a_{GH} z^H$ if $A = (a_{GH})$. Then

$$z^T = A^{-1}(w^T) = A^{-1}S_T(w) = A^{-1}S_T A(z).$$

The condition for A is therefore $A^{-1}S_T A = S_T$, $|A| \neq 0$. The most general

matrix commutative with every S_T is a matrix of the second regular representation. If α is an element of the group ring, this representation is defined by

$$\alpha G = \Sigma a_{GH} H.$$

It is a reciprocal representation (A' is a direct representation). Then $w = A(z)$ or $w^G = \Sigma a_{GH} z^H$ reads $w^G = z^{\alpha G}$. The element α must not be a divisor of zero because otherwise $|A| = 0$.

BRYN MAWR COLLEGE.

BIBLIOGRAPHY.

1. R. Brauer, "Über die Kleinsche Theorie der algebraischen Gleichungen," *Mathematische Annalen*, vol. 110 (1934), § 2, pp. 482.

3. M. Deuring, "Algebren," *Ergebnisse der Mathematik*, vol. 4 (1935).

4. M. Deuring, "Galoissche Theorie und Darstellungstheorie," *Mathematische Annalen*, vol. 107 (1933), pp. 140.

5. E. L. Dickson, *Algebras and their Arithmetics*, Chicago, 1923.

6. H. Hasse, *Klassenkörpertheorie*, 1932-33, p. 157.

7. K. Hensel, "Über die Darstellung der Zahlen eines Gattungsbereiches für einen beliebigen Primdivisor," *Journal für die reine und angewandte Mathematik*, vol. 103 (1888), pp. 230.

8. E. Noether, "Hyperkomplexe Grössen und Darstellungstheorie," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 642-692.

9. E. Noether, "Nichtkommutative Algebra," *Mathematische Zeitschrift*, vol. 37 (1933), p. 538.

10. E. Noether, "Normalbasis bei Körpern ohne höhere Verzweigung," *Journal für die reine und angewandte Mathematik*, vol. 167 (1931-32), p. 149.

11. A. Speiser, *Gruppen theorie*, J. Springer, Berlin, 1927.

12. A. Speiser, "Gruppendeterminante und Körperdiskriminante," *Mathematische Annalen;* vol. 77 (1916),1, p. 546; 2, p. 552.

13. B. L. Van der Waerden, *Moderne Algebra*, J. Springer, Berlin, 1931.

14. J. H. M. Wedderburn, "Lectures on Matrices," *American Mathematical Society Colloquium Publications*, New York, 1934.

A REMARK ON THE AREA OF SURFACES.[1]

By Tibor Radó.

Introduction. Let Σ be a continuous surface given by equations

$$\Sigma: \; x = x(u,v), \quad y = y(u,v), \quad z = z(u,v), \quad 0 \leqq u \leqq 1, \quad 0 \leqq v \leqq 1.$$

One of the fundamental problems in the theory of the area is to find conditions under which the Lebesgue area $L[\Sigma]$ of Σ is given by the classical formula

$$L[\Sigma] = \int_0^1 \int_0^1 (EG - F^2)^{1/2} du dv,$$

where E, F, G have the familiar meaning.[2] In the special case $x = u$, $y = v$ Tonelli found that the absolute continuity, in the sense defined by him, of the function $z(u,v)$ is a necessary and sufficient condition for the validity of the classical formula for the area.[3] In the general case, only sufficient conditions were established so far. The most general result in this direction was obtained by McShane and by Morrey.[4] To make our remarks more definite, we shall consider the work of Morrey. Morrey defines a class L of surfaces by requiring the existence of representations where the coördinate functions $x(u,v)$, $y(u,v)$, $z(u,v)$ satisfy two conditions (i) and (ii).[5] Condition (ii) is concerned with the approximation of the coördinate functions by integral means. Condition (i) requires that the coördinate functions be absolutely continuous in the sense of Tonelli. Both McShane and Morrey observe that in the special case $x = u$, $y = v$ these conditions are equivalent to the necessary and sufficient condition of Tonelli. It might be therefore of some interest to point out that condition (i) can be replaced by the weaker condition that the coördinate

[1] Presented to the American Mathematical Society at the meeting in Chicago, April, 1936.

[2] For general information and for references to the literature concerning continuous surfaces and the Lebesgue area the reader may consult the paper of C. B. Morrey, "A class of representations of manifolds, part I," *American Journal of Mathematics*, vol. 55 (1933), pp. 683-707.

[3] L. Tonelli, "Sulla quadratura delle superficie," *Atti della Reale Accademia dei Lincei*, series 6, vol. 3 (1926), pp. 357-362, 445-450, 633-638, 714-719. S. Saks gave a very elegant presentation of the results of Tonelli in his paper "Sur l'aire des surfaces $z = f(x,y)$," *Acta Szeged*, vol. 3 (1927), pp. 170-176.

[4] E. J. McShane, "Integrals over surfaces in parametric form," *Annals of Mathematics*, vol. 34 (1933), pp. 815-838. C. B. Morrey, *loc. cit.* [2].

[5] *Loc. cit.* [2], p. 701.

functions be of bounded variation in the sense of Tonelli. If we denote this weaker condition by (i*), and if we use $\mathfrak{K}$ to denote the class of surfaces defined by conditions (i*) and (ii), then our main result states that the classical formula for the Lebesgue area $L[\Sigma]$ holds for every surface Σ of class $\mathfrak{K}$ (provided, of course, that we use a representation satisfying conditions (i*) and (ii)). In the special case $x=u$, $y=v$ conditions (i*) and (ii) are still equivalent to the necessary and sufficient condition of Tonelli. Our argument will be based upon certain simple facts concerning the topological index which were established in a previous paper of the author [6] and which we are going to state presently.

Let C_k be a sequence of closed continuous curves, in the (x,y)-plane, which converge in the sense of Fréchet to a closed continuous curve C. This means that these curves admit of simultaneous representations

$$C:\ x=x(t),\quad y=y(t),\quad 0\leqq t\leqq 1,\quad x(0)=x(1),\quad y(0)=y(1),$$
$$C_k:\ x=x_k(t),\quad y=y_k(t),\quad 0\leqq t\leqq 1,\quad x_k(0)=x_k(1),\quad y_k(0)=y_k(1),$$

where the functions $x(t)$, $y(t)$, $x_k(t)$, $y_k(t)$ are continuous in $0\leqq t\leqq 1$ and $x_k(t)\rightarrow x(t)$, $y_k(t)\rightarrow y(t)$ uniformly in $0\leqq t\leqq 1$. Let $n(x,y)$, $n_k(x,y)$ be the index-functions relative to C, C_k respective (see 2.3). Let us use $T[\lambda]$ to denote the total variation of a function $\lambda(t)$ defined in $0\leqq t\leqq 1$, where $T[\lambda]=\infty$ if $\lambda(t)$ is not of bounded variation. It is well known that if at least one of the functions $x(t)$, $y(t)$ is of bounded variation, then the index-function $n(x,y)$ is summable.[7] Using these notations, we have the following lemmas.

LEMMA 1. *If $C_k\rightarrow C$ in the sense of Fréchet, and if $T[x_k]\rightarrow T[x]<\infty$ (or $T[y_k]\rightarrow T[y]<\infty$), then*

$$\int\int|\,n(x,y)-n_k(x,y)\,|\,dxdy\rightarrow 0.$$

LEMMA 2. *If 1) $C_k\rightarrow C$ in the sense of Fréchet, 2) the functions $x(t),y(t),x_k(t),y_k(t)$ are of bounded variation, 3) $T[x_k]<M$, $T[y_k]<M$, where M is some finite constant independent of k, then*

$$\int\int n_k(x,y)\,dxdy\rightarrow\int\int n(x,y)\,dxdy.$$

[6] "A lemma on the topological index," submitted for publication to the editors of the *Fundamenta Mathematicae*. An abstract appeared in the *Bulletin of the American Mathematical Society*, vol. 42 (1936), p. 187.

[7] J. Schauder, "Über stetige Abbildungen," *Fundamenta Mathematicae*, vol. 12 (1928), pp. 47-74, in particular pp. 64-66.

Lemma 2 is a consequence of lemma 1, but can be proved independently also. The reader will note that the conclusion is much weaker in lemma 2 than in lemma 1. In the situation which we shall consider, the assumptions of the strong lemma 1 will be amply satisfied, while we shall actually need only the conclusion of the weak lemma 2. In a general way, despite the generality of our main theorem it will be quite apparent that the definition of the class $\mathfrak{K}$ implies considerably more than what is actually needed in the proof. With regard to possible use in investigations suggested by this remark, we develop certain inferences from our assumptions beyond the absolute minimum needed for our present purposes. We also took the liberty of stating condition (ii) of Morrey in an obviously equivalent but somewhat more convenient form, by splitting it into conditions II and III of section 2. 1.

INDEX OF NOTATIONS AND DEFINITIONS.

1. *On the approximation by integral means.*

1. 1. Let $x(u, v)$ be a continuous function in the closed square S_0: $0 \leqq u \leqq 1$, $0 \leqq v \leqq 1$. For $0 < h < \frac{1}{2}$, the closed square $h \leqq u \leqq 1 - h$, $h \leqq v \leqq 1 - h$ will be denoted by S_h. In S_h, we define

$$x^{(h)}(u, v) = (1/4h^2) \int_{-h}^{h} \int_{-h}^{h} x(u + \alpha, v + \beta)\, d\alpha d\beta.$$

As it is well known, $x^{(h)}(u, v)$ is continuous in S_h together with its partial derivatives of the first order.[8]

[8] For a systematic presentation of the properties of this approximation see H. E. Bray, "Proof of a formula for an area," *Bulletin of the American Mathematical Society*, vol. 29 (1923), pp. 264-270.

1. 2. Let R: $a \leqq u \leqq b$, $c \leqq v \leqq d$ be a rectangle completely interior to S_0. For small values of h, R will be completely interior to S_h also and clearly $x^{(h)}(u,v) \to x(u,v)$ for $h \to 0$ uniformly in R.

1. 3. Suppose that $x(u,v)$, considered as a function of v alone, is of bounded variation on the interval $u = \text{const.}$, $v_1 \leqq v \leqq v_2$. The total variation on this interval will be denoted by ${}_vT_{v_1}{}^{v_2}[u;x]$. If $x(u,v)$ is not of bounded variation on this interval, we put ${}_vT_{v_1}{}^{v_2}[u;x] = \infty$. The symbols ${}_uT_{u_1}{}^{u_2}[v;x]$, ${}_vT_{v_1}{}^{v_2}[u;x^{(h)}]$, ${}_uT_{u_1}{}^{u_2}[v;x^{(h)}]$ are defined in a similar fashion.

1. 4. Let R: $a \leqq u \leqq b$, $c \leqq v \leqq d$ be a rectangle completely interior to S_0. We define

$$T[R;x] = {}_uT_a{}^b[c;x] + {}_vT_c{}^d[b;x] + {}_uT_a{}^b[d;x] + {}_vT_c{}^d[a;x]$$

if all the terms on the right-hand side are finite. Otherwise we put $T[R;x] = \infty$. For small values of h, R will be completely interior to S_h also, and we define then the symbol $T[R;x^{(h)}]$ in a similar way.

1. 5. The continuous function $x(u,v)$ is of bounded variation in the sense of Tonelli if ${}_vT_0{}^1[u;x]$, ${}_uT_0{}^1[v;x]$ are summable in the intervals $0 \leqq u \leqq 1$, $0 \leqq v \leqq 1$ respectively. To describe this situation, we shall say that $x(u,v)$ is B. V. T. in S_0.

1. 6. Let the continuous function $x(u,v)$ be B. V. T. in S_0. The function ${}_vT_0{}^1[u;x]$ being then summable, the function ${}_vT_{v_1}{}^{v_2}[u;x]$, where $0 \leqq v_1 < v_2 \leqq 1$, is *a fortiori* summable in the interval $0 \leqq u \leqq 1$ for fixed v_1, v_2. By a well-known theorem of the Lebesgue theory, we have therefore in the interval $0 \leqq u \leqq 1$ a set ${}_uE[v_1,v_2;x]$ of measure zero, such that ${}_vT_{v_1}{}^{v_2}[u;x]$ is finite and

$$(1/2h)\int_{-h}^{h} {}_vT_{v_1}{}^{v_2}[u+\alpha;x]\,d\alpha \underset{h\to 0}{\longrightarrow} {}_vT_{v_1}{}^{v_2}[u;x]$$

for u not in ${}_uE[v_1,v_2;x]$. Similarly, we have in the interval $0 \leqq v \leqq 1$ a set ${}_vE[u_1,u_2;x]$ of measure zero such that ${}_uT_{u_1}{}^{u_2}[v;x]$ is finite and

$$(1/2h)\int_{-h}^{h} {}_uT_{u_1}{}^{u_2}[v+\beta;x]\,d\beta \underset{h\to 0}{\longrightarrow} {}_uT_{u_1}{}^{u_2}[v;x]$$

for v not in ${}_vE[u_1,u_2;x]$.

1. 7. We denote by ${}_uE[x]$ the sum of all the sets ${}_uE[v_1,v_2;x]$ which correspond to all pairs of *rational* numbers v_1, v_2 such that $0 \leqq v_1 < v_2 \leqq 1$. The set ${}_vE[x]$ is defined in a similar way. The sets ${}_uE[x]$, ${}_vE[x]$ are both

sums of a denumerable infinity of sets of measure zero, and are therefore of measure zero.

1.8. A rectangle R: $a \leqq u \leqq b$, $c \leqq v \leqq d$ will be called *admissible* with respect to $x(u, v)$ if a, b are not in ${}_uE[x]$, c, d are not in ${}_vE[x]$, and if R is completely interior to S_0.

1.9. Lemma. *If the continuous function $x(u, v)$ is B. V. T. in S_0, then*

$$T[R; x^{(h)}] \underset{h\to 0}{\to} T[R; x]$$

for every admissible rectangle R.[9]

To prove this, it is sufficient to discuss any one of the four terms which make up $T[R; x]$. Let us show, for instance, that

$$(1) \qquad {}_uT_a{}^b[c; x^{(h)}] \underset{h\to 0}{\to} {}_uT_a{}^b[c; x].$$

Since the interval $v = c$, $a \leqq u \leqq b$ is completely interior to S_0, we have $x^{(h)} \to x$ uniformly on this interval. Hence obviously

$$(2) \qquad \varliminf_{h\to 0} {}_uT_a{}^b[c; x^{(h)}] \geqq {}_uT_a{}^b[c; x].$$

Let now u', u'' be any two *rational* numbers such that

$$(3) \qquad 0 < u' < a < b < u'' < 1.$$

Let us take any system of numbers

$$u_0 = a < u_1 < \cdots < u_{k-1} < u_k < \cdots < u_n = b.$$

We have then, for $h < u'' - b$, $h < a - u'$,

$$\begin{aligned}
&\sum_{k=1}^{n} | x^{(h)}(u_k, c) - x^{(h)}(u_{k-1}, c) | \\
&\quad \leqq (1/4h^2) \int_{-h}^{h} \int_{-h}^{h} \sum_{k=1}^{n} | x(u_k + \alpha, c + \beta) - x(u_{k-1} + \alpha, c + \beta) | \, d\alpha d\beta \\
&\quad \leqq (1/4h^2) \int_{-h}^{h} \int_{-h}^{h} {}_uT_{a+\alpha}^{b+\alpha}[c + \beta; x] \, d\alpha d\beta \\
&\quad \leqq (1/4h^2) \int_{-h}^{h} \int_{-h}^{h} {}_uT_{u'}{}^{u''}[c + \beta; x] \, d\alpha d\beta = (1/2h) \int_{-h}^{h} {}_uT_{u'}{}^{u''}[c + \beta; x] \, d\beta.
\end{aligned}$$

---

[9] The idea of the proof was suggested by a similar argument used by S. Saks, *loc. cit.* [8]. Morrey, *loc. cit.* [2], p. 703, shows (in our notations) that

$$\varlimsup_{h\to 0} T[R; x^{(h)}] < \infty$$

by a reasoning which makes use of absolute continuity of the functions involved.

Consequently

$$(4) \qquad {}_uT_a{}^b[c;x^{(h)}] \leqq (1/2h)\int_{-h}^{h} {}_uT_{u'}{}^{u''}[c+\beta;x]\,d\beta.$$

Since u', u'' are rational and c is not in ${}_vE[x]$, we infer from (4) that

$$(5) \qquad \overline{\lim_{h\to 0}}\, {}_uT_a{}^b[c;x^{(h)}] \leqq {}_uT_{u'}{}^{u''}[c;x].$$

But $x(u,v)$ is continuous and consequently, for fixed c, ${}_uT_{u'}{}^{u''}[c;x]$ is a continuous function of u', u''. Hence (5) yields, for $u' \to a$, $u'' \to b$, the relation

$$(6) \qquad \overline{\lim_{h\to 0}}\, {}_uT_a{}^b[c;x^{(h)}] \leqq {}_uT_a{}^b[c;x].$$

(2) and (6) imply (1), and the proof is complete.

2. *On transformations of class* $\mathfrak{K}$.

2.1. If $x(u,v)$, $y(u,v)$ are continuous in S_0, then the equations $x = x(u,v)$, $y = y(u,v)$ define a continuous transformation. We shall say that this transformation is of class $\mathfrak{K}$ if the following conditions are satisfied.[10]

I. $x(u,v)$, $y(u,v)$ are B. V. T. in S_0.

II. The Jacobian [11]

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

is summable in S_0.

III. For every rectangle R: $a \leqq u \leqq b$, $c \leqq v \leqq d$, which is completely interior to S_0, we have

$$\iint_R \left| \frac{\partial(x,y)}{\partial(u,v)} - \frac{\partial(x^{(h)},y^{(h)})}{\partial(u,v)} \right| du\,dv \underset{h\to 0}{\longrightarrow} 0.$$

In the last relation, $y^{(h)}$ is defined in the same way in terms of y as $x^{(h)}$ was defined in terms of x.

2.2. A rectangle R: $a \leqq u \leqq b$, $c \leqq v \leqq d$ will be called admissible with respect to the transformation if it is admissible with respect to both $x(u,v)$ and $y(u,v)$, in the sense of 1.8.

---

[10] Conditions II and III are obviously equivalent to condition (ii) of Morrey, while condition I replaces the more restrictive condition (i) of Morrey which requires absolute continuity in the sense of Tonelli. See Morrey, *loc. cit.* [2], p. 701.

[11] On account of condition I, the partial derivatives of the first order exist almost everywhere in S_0.

2.3. Under the continuous transformation $x = x(u, v)$, $y = y(u, v)$ the boundary B of a rectangle R (comprised in S_0) is carried into a closed continuous curve C in the (x, y)-plane. If (x, y) is a point not on C, we define $n_R(x, y)$ as the topological index [12] of (x, y) with respect to C, this curve being described in the sense which corresponds to the counter-clockwise sense around R. If (x, y) is on C, we put $n_R(x, y) = 0$.

2.4. If the rectangle R is completely interior to S_0, then $x^{(h)}$, $y^{(h)}$ will be both defined on R for small values of h. The symbol ${}_hn_R(x, y)$ is then defined in the same way in terms of $x^{(h)}(u, v)$, $y^{(h)}(u, v)$ as $n_R(x, y)$ was defined in terms of $x(u, v)$, $y(u, v)$.

2.5. LEMMA. *If the continuous transformation $x = x(u, v)$, $y = y(u, v)$ is of class $\mathfrak{K}$ in S_0, then*

$$\int\int | n_R(x, y) - {}_hn_R(x, y) | \, dxdy \underset{h \to 0}{\longrightarrow} 0 \tag{7}$$

and

$$\int_R\int \frac{\partial(x, y)}{\partial(u, v)} \, dudv = \int\int n_R(x, y) \, dxdy \tag{8}$$

for every admissible rectangle R.[13]

Proof. We have $x^{(h)} \to x$, $y^{(h)} \to y$ uniformly on the boundary of R, and since R is admissible, we have, by 1.9,

$$T[R; x^{(h)}] \underset{h \to 0}{\longrightarrow} T[R; x], \tag{9}$$

$$T[R; y^{(h)}] \underset{h \to 0}{\longrightarrow} T[R; y]. \tag{10}$$

The relation (7) is thus a direct consequence of lemma 1 (see the introduction). Since $x^{(h)}$, $y^{(h)}$ have continuous derivatives of the first order, we have [14]

$$\int_R\int \frac{\partial(x^{(h)}, y^{(h)})}{\partial(u, v)} = \int\int {}_hn_R(x, y) \, dxdy. \tag{11}$$

---

[12] See, for instance, Kerékjártó, *Vorlesungen über Topologie*, vol. 1, section 2, § 2.

[13] Formula (8) is a generalization of Lemma 4 of Morrey, *loc. cit.* [2], p. 702. For the case when $x(u, v)$, $y(u, v)$ satisfy the Lipschitz condition, formula (8) was established by Schauder, *loc. cit.* [7]. Formula (7) is independent of conditions II and III in 2.1 and expresses a property of the approximation by integral means which apparently was not yet noticed.

[14] See Schauder, *loc. cit.* [7].

Condition III in 2. 1 implies that

$$(12)\qquad \lim_{h\to 0}\iint_R \frac{\partial(x^{(h)}, y^{(h)})}{\partial(u, v)}\,dudv = \iint_R \frac{\partial(x, y)}{\partial(u, v)}\,dudv,$$

while (7) implies that

$$(13)\qquad \lim_{h\to 0}\iint {}_h n_R(x, y)\,dxdy = \iint n_R(x, y)\,dxdy.$$

Clearly, (11), (12) and (13) imply (8).

3. *Conclusion.*

3. 1. We shall say that a continuous surface Σ is of class $\mathfrak{K}$ if it admits of a representation

$$\Sigma:\ x = x(u, v),\quad y = y(u, v),\quad z = z(u, v),\quad (u, v) \text{ in } S_0,$$

such that each of the three transformations

$$\begin{aligned} y &= y(u, v), & z &= z(u, v),\\ z &= z(u, v), & x &= x(u, v),\\ x &= x(u, v), & y &= y(u, v) \end{aligned}$$

is of class $\mathfrak{K}$ in S_0, in the sense of 2. 1. Every representation with this property will be called a typical representation of Σ.

3. 2. If in the definition of transformations of class $\mathfrak{K}$ we replace condition I (see 2. 1) by the more restrictive condition that the defining functions be absolutely continuous in the sense of Tonelli, and if we modify the definition of surfaces of class $\mathfrak{K}$ accordingly, then we obtain the surfaces of class L studied by McShane and Morrey.[15] To make the following remarks more concise, we shall consider the work of Morrey. The assumption of absolute continuity is used by Morrey only to establish formula (8) (see 2. 5). We derived that formula under the assumption of bounded variation in the sense of Tonelli. This being so, it is clear that we can proceed in exactly the same way in the case of surfaces of class $\mathfrak{K}$ as Morrey did in the case of surfaces of class L.[16] As a result, we obtain the following

[15] The remarks of McShane, *loc. cit.* [4] and of Morrey, *loc. cit.* [2], concerning the generality of the class L apply *a fortiori* to the class $\mathfrak{K}$.

[16] We are referring to the proof of Theorem I of Morrey, *loc. cit.* 2, pp. 703-704.

THEOREM. *If Σ is a surface of class $\mathfrak{K}$ given in typical representation, then*

$$L[\Sigma] = \iint_{S_0} (EG - F^2)^{1/2} du dv, \tag{14}$$

where $L[\Sigma]$ is the Lebesgue area of Σ and

$$E = x_u^2 + y_u^2 + z_u^2, \quad F = x_u x_v + y_u y_v + z_u z_v, \quad G = x_v^2 + y_v^2 + z_v^2.$$

3. 3. We conclude with a few remarks to the effect that even the weaker assumption of bounded variation (in the sense of Tonelli) implies considerably more than what we actually need to derive (14). Inspection of the proof shows that we needed bounded variation only to establish formula (8) in 2. 5. To establish that formula it would be sufficient to know that

$$\iint {}_h n_R(x, y)\, dx dy \underset{h\to 0}{\longrightarrow} \iint n_R(x, y)\, dx dy, \tag{15}$$

while bounded variation implies (see 2. 5) the much stronger relation

$$\iint |\, n_R(x, y) - {}_h n_R(x, y)\, |\, dx dy \underset{h\to 0}{\longrightarrow} 0. \tag{16}$$

In order to establish (15) it would be sufficient to know that

$$\overline{\lim_{h\to 0}}\, T[R; x^{(h)}] < \infty, \quad \overline{\lim_{h\to 0}}\, T[R; y^{(h)}] < \infty, \tag{17}$$

as it follows from lemma 2 (see the introduction), while bounded variation yields (see 1. 9) the more precise information

$$\lim_{h\to 0} T[R; x^{(h)}] = T[R; x], \quad \lim_{h\to 0} T[R; y^{(h)}] = T[R; y]. \tag{18}$$

It follows further from lemma 1 (see the introduction) that only one of the two relations (18) is needed to establish formula (8) in 2. 5. These remarks suggest the possibility of further generalizations.

THE OHIO STATE UNIVERSITY.

ON THE POINCARÉ GROUP OF RATIONAL PLANE CURVES.

By Oscar Zariski.

Introduction. In a paper dealing with the Poincaré group of an algebraic hypersurface V_{n-1} in a projective complex space S_n (Zariski [9]), the following theorem is proved: *the Poincaré group of the residual space $S_n - V_{n-1}$ coincides with the Poincaré group of the residual space of a generic plane section of V_{n-1}.* In this paper we apply this theorem toward the determination of the Poincaré group of any rational plane curve with nodes and cusps only, and, more generally, of any plane curve which admits such a rational curve as a limiting case. It would seem that the sense of the quoted theorem in applications would be to reduce the apparently more difficult problem of the Poincaré group of an hypersurface to that of the Poincaré group of a plane algebraic curve. However, in this paper we apply the theorem in the opposite sense, using a convenient hypersurface in order to solve the problem in the plane. The advantage of transforming the plane problem into a problem in a space of higher dimension seems due to the fact (at least it is so in the present case) that if a curve C is a generic plane section of V_{m-1}, then, everything else being equal, the hypersurface V_{n-1} supplies a more intrinsic picture of the Poincaré group of C than C itself. The essential features of the Poincaré group of C should be revealed best on the hypersurface V_{n-1} lying in a space of the highest possible dimension.

In the present paper the starting point is supplied by the class of rational *maximal cuspidal curves* of even order $2n-2$. These curves are generic plane sections of what may be referred to as the discriminant hypersurface D: $D(a_0, \cdots, a_n) = 0$, where D is the discriminant of the polynomial $a_0 z^n + \cdots + a_n$, and where $a_0, a_1, \cdots, a_n$ are interpreted as homogeneous coördinates in an S_n. The corresponding class of Poincaré groups G_n practically coincides with the "Zopfgruppe" of Artin (Artin,[1] see also Reidemeister,[4] p. 42). G_n is also the group of automorphism classes of a sphere with n holes (Magnus [3]). We give a proof of the completeness of the generating relations of G_n which is simpler than the previous proofs.

The case of an arbitrary rational curve with nodes and cusps is treated by applying the notion of virtually non-existent nodes or cusps (Severi,[6] Anhang F, B. Segre,[5] Zariski,[7] p. 168), and by studying the effect on the

Poincaré group of a plane curve of the removal of a cusp or of a node. This can be done, since the type of generating relations at a cusp or at a node is known (Zariski [8]). The result is largely negative: *every rational curve with nodes and cusps, other than the maximal cuspidal curve of even order* $2n-2$, *has, with one exception* (see Section 6), *a cyclic Poincaré group. The same is true of any plane algebraic curve which admits a rational curve with nodes and cusps as a limiting case.*

1. *Preliminary remarks on continuous systems of rational curves.* A rational curve with order n with k cusps shall be denoted by the symbol (n,k). Applying the formulae of Plücker, it is seen that the dual of a curve (n,k) is a curve (n',k'), where

$$(1) \qquad n'=2n-2-k, \qquad k'=3(n-2)-2k.$$

Hence $k\leqq \frac{3}{2}(n-2)$. If n is even, the rational maximal cuspidal curves $(n, \frac{3}{2}(n-2))$ are dual to the curves $((n+2)/2, 0)$ (rational curves possessing only nodes). If n is odd, the maximal cuspidal curves $(n, (3n-7)/2)$ are dual to the curves $((n+3)/2, 1)$. From this it follows that *the maximal cuspidal curves form in either case a single irreducible continuous system* (Severi,[6] Anhang F). Since the characteristic series of this system is non-special, any number of cusps of the general curve of the system can be converted into nodes (*virtual nodes*), *and hence curves* (n,k) *exist for every integral value of* k *satisfying the inequality* $k\leqq 3(n-2)/2$ (see, for instance, B. Segre [5]).

For any given k, *satisfying the above inequality, the curves* (n,k) *form a single irreducible continuous system.* For the proof we observe that it is indifferent whether the statement is proved for the curves (n,k) or for the dual curves (n',k'). From (1) follows the relation $(n-2)-k=k'-(n'-2)$, and hence either $n-2-k$ or $n'-2-k'$ is a non-negative integer. We may assume therefore $k\leqq n-2$. A curve (n,k) is the projection of a normal rational curve $\Gamma_n{}^n$ of order n, in S_n, the center of projection being an S_{n-3} meeting in k points the ruled surface F of the tangents of $\Gamma_n{}^n$. If $k\leqq n-2$, any k points belong to an S_{n-3} and hence the k intersections of S_{n-3} with F can be taken arbitrarily on F. Our statement now follows from the irreducibility of the system of k-ads of points of F (F is irreducible!) and from the irreducibility of the system of all S_{n-3}'s on k fixed points.

Since *a* system of curves (n,k) can be obtained from the system of maximal cuspidal curves by regarding a certain number of cusps of the general curve of this system as virtual nodes, we conclude that the complete system of curves

(n, k) contains for any $k \leqq \frac{3}{2}(n-2)$, the complete system of rational maximal cuspidal curves. In other words, *any rational cuspidal curve of order n possesses the maximal cuspidal curve of order n as a limit case.*

2. *The Poincaré group of the maximal cuspidal curve of even order.* Let C_{2n-2} be a maximal cuspidal curve $(2n-2, 3(n-2))$, of even order $2n-2$, the dual of a rational plane curve Γ_n of order n, possessing nodes only. For the general curve Γ_n we have the parametric equations:

$$x_i = f_i(t), \qquad (i = 1, 2, 3), \tag{2}$$

where f_1, f_2, f_3 are arbitrary polynomials of degree n in t. The equation of C_{2n-2} is $D(\lambda_1, \lambda_2, \lambda_3) = 0$, where $D(\lambda_1, \lambda_2, \lambda_3)$ is the discriminant of the polynomial $F(t) = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$. If we interpret the coefficients a_i of the general polynomial $f(t) = a_0 t^n + a_1 t^{n-1} + \cdots + a_n$ of degree n as homogeneous coördinates of a point in a complex projective space S_n, we see that C_{2n-2} is the intersection of the plane $F = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$ with the hypersurface $D(a_0, a_1, \cdots, a_n) = 0$, where D is the discriminant of $f(t)$. We denote this hypersurface by Δ and we shall refer to Δ as the *discriminant hypersurface.* By the theorem on the Poincaré group of an algebraic hypersurface, quoted in the introduction, we have that the Poincaré group of C_{2n-2} coincides with the Poincaré group of Δ (i. e. of the residual space $S_n - \Delta$). We shall denote this group by G_n, and by g an element of G_n.

We interpret G_n as *the group of motions of n distinct points on the sphere H of the complex variable t.* Each point f $(= f(t))$ of S_n represents an unordered set of n points $t_1, t_2, \cdots, t_n$ of H, the roots of the polynomial $f(t)$. If f is on $S_n - \Delta$, these n points are distinct, and conversely. A closed path in $S_n - \Delta$ corresponds to a motion g of the points t_i on H, in the course of which these n points remain always distinct from each other and which carries a given unordered set $(t_1{}^0, t_2{}^0, \cdots, t_n{}^0)$ into its initial position. A slight deformation of g over $S_n - \Delta$ corresponds to a slight deformation of the given motion g, consisting both in a deformation of the paths and, so to speak, of the instantaneous velocities of the individual points t_i, the variable set $(t_1, t_2, \cdots, t_n)$ consisting always of n distinct points in the course of the deformation. If $g = 1$, the given motion can be deformed in the above manner into rest, the initial set $(t_1{}^0, \cdots, t_n{}^0)$ remaining fixed during the deformation.

The consideration of the group G_n goes back to Hurwitz [2] who has applied it toward the classification of Riemann surfaces with assigned branch points. In this paper Hurwitz determines the generators of G_n. The group G_n,

interpreted as the group of automorphism classes of the sphere H with n holes (the holes being at the points $t_i{}^0$ of the initial set), has been studied by Magnus,[3] who derived the generating relations of G_n and who pointed out the connection between the groups G_n and the "Zopfgruppen" of Artin.[1] In view of the importance of this class of groups, we give here a new and simpler treatment of the group G_n. We point out that the $n-1$ generating relations given in the quoted paper of Magnus (Magnus,[3] relations (19)) are all consequences of the one relation (6) given below.

3. *The generators of G_n.* We fix on H a set of n distinct points $P_1, P_2, \cdots, P_n$ as an initial set and we denote by $X_1, X_2, \cdots, X_n$ the points of a variable set. We join the points $P_1, P_2, \cdots, P_n$, in the order written, by a simple oriented arc, and we denote by s_i the oriented arc joining the points P_i, P_{i+1}. Let g_i denote the motion in which the points P_j, $j \neq i, i+1$, are fixed, while the points P_i and P_{i+1} are interchanged, X_i moving from P_i to P_{i+1} along the right-hand edge of the oriented arc s_i and X_{i+1} moving from P_{i+1} toward P_i along the opposite edge of s_i (Fig. 1). We prove that *the $n-1$ elements $g_1, g_2, \cdots, g_{n-1}$ are generators of G_n.*

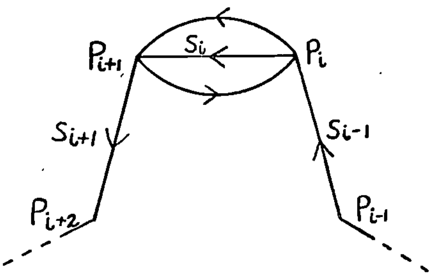

FIG. 1.

Proof. The elements g_i, considered as transpositions $(P_i\,P_{i+1})$, generate the symmetric group of permutations of the n points P_i. Hence, every element of G_n can be written as a product of g_i's multiplied by an element S, representing a motion in which each point X_i comes back to its original position P_i. Let S also denote the corresponding singular 1-sphere in the residual space $S_n - \Delta$; its initial point is $P \equiv (P_1, P_2, \cdots, P_n)$, while a variable point of S_n representing a variable set of n points $(X_1, X_2, \cdots, X_n)$ shall be denoted by X. If Q is a fixed point on the sphere H, we denote by V_Q the $(n-1)$-dimensional variety in S_n representing the sets of n points of H containing Q. We first deform S in such a manner that the closed path described by the point X_1, starting from and returning to P_1, does not pass through the points $P_2, P_3, \cdots, P_n$. We join then each point $X \equiv (X_1, X_2, \cdots, X_n)$ of S to the point $X' \equiv (X_1, P_2, \cdots, P_n)$ by a simple arc l contained in V_{X_1} but not meeting Δ. As X varies on S, we vary the arc l continuously, assuming that when X is very near its initial position P (and hence X' is very near X), the arc l is very small and reduces to a point at the initial position P of X.

The final position of l, as X describes the entire 1-sphere S, is a singular 1-sphere S' on V_{P_1}, in the residual space of Δ, while the point X' describes a singular 1-sphere S_1. The locus of the arc l is a singular 2-cell bounded by $S(S_1S')^{-1}$. Hence we can deform S into the product S_1S', where S_1 represents a motion in which all the points P_i, except P_1, are fixed, and S' is a motion in which the point P_1 is fixed. The same procedure can now be applied to S' and yields a deformation of S' into a product S_2S'', where the motion S_2 leaves all the points P_i, except P_2, fixed and where in S'' the points P_1 and P_2 are fixed. Continuing in this manner we finally express S as a product $S_1S_2\cdots S_n$, where S_α is a motion in which all the points P_i, except P_α, are fixed. Now S_α is a singular 1-sphere in $H - P_1 - P_2 - \cdots - P_{\alpha-1} - P_{\alpha+1} - \cdots - P_n$ and can be deformed into a product of loops issued from the point P_α and surrounding the points $P_1, \cdots, P_{\alpha-1}, P_{\alpha+1}, \cdots, P_n$. It is easily seen that such loops are supplied by products of the elements g_i. For instance, if $\alpha = 1$, then the products $g_1^2, g_1^{-1}g_2^2g_1, \cdots, (g_{n-2}\cdots g_1)^{-1}\, g^2_{n-1}(g_{n-2}\cdots g_1)$ represent the required loops,[1] q. e. d.

4. *The generating relations of* G_n. We have in the first place the following relations between the generators g_i:

$$g_ig_j = g_jg_i, \qquad |i-j| \neq 1. \tag{3}$$

[1] As originally defined, g_1^2 is a motion in which the variable point X_1, starting from P_1, turns about P_2 in the positive (counterclockwise) sense, while *at the same time* the point X_2, starting from P_2, turns about P_1 in the same sense. However, as can be seen from the accompanying figure (Fig. 2), the loop described by either one of these points, say by X_2, can be pulled over the point P_1 and then deformed into the point P_2, while the path of X_1 is unaltered.

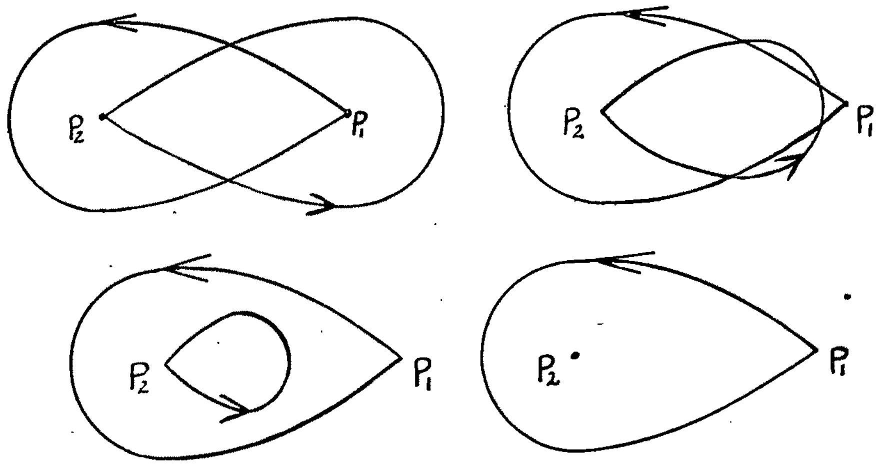

FIG. 2.

These relations are trivial, because if $|i-j|\neq 1$, the arcs s_i and s_j have no end-points in common and it is therefore indifferent which one of the two motions g_i and g_j takes place first.

We now consider an oriented arc s'_i joining the points P_i and P_{i+2}, very near but not meeting the arc s_i+s_{i+1} outside the end points. We assume that s'_i is on the left-hand side of the oriented arc s_i+s_{i+1} and that P_{i+2} is its initial point. Let g'_i be the motion in which the points P_i and P_{i+2} are interchanged along the arc s'_i, the point X_{i+2} moving from P_{i+2} toward P_i on the right-hand edge of s'_i. It is then easily seen that the motion in which the three points P_i, P_{i+1}, P_{i+2} are permuted cyclically, their paths being the oriented arcs s_i, s_{i+1}, s'_i respectively (Fig. 3) is expressable in terms of the elements g_i, g'_i, g_{i+1} in each of the three following manners: $g_{i+1}g_i$, g'_ig_{i+1}, $g_ig'_i$. Hence $g_{i+1}g_i=g_i'g_{i+1}=g_ig'_i$, and eliminating g'_i, we find the following relation:

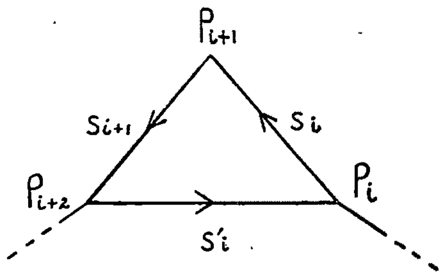

FIG. 3.

$$g_ig_{i+1}g_i=g_{i+1}g_ig_{i+1}, \qquad (i=1,2,\cdots,n-2). \tag{4}$$

We have pointed out in the preceding section that loops, which we shall denote by $a_2, a_3,\cdots,a_n$, which issue from the point P_1 and surround the points $P_2, P_3,\cdots,P_n$ respectively, are given by the following products:

$$a_i=(g_{i-2}\cdots g_1)^{-1}g^2_{i-1}(g_{i-2}\cdots g_1). \tag{5}$$

Since the product $a_2a_3\cdots a_n$ is obviously the identity in G_n, we have

$$g_1g_2\cdots g_{n-2}g^2_{n-1}g_{n-2}\cdots g_2g_1=1. \tag{6}$$

We proceed to prove that *the relations* (3), (4) *and* (6) *constitute a complete set of generating relations of* G_n.

We introduce the following notation. If two power products Πg_i and $\Pi' g_i$ of $g_1,\cdots,g_n$ are equal *as a consequence of the relations* (3), (4) *and* (6) *only*, then we shall write $\Pi g_i\equiv\Pi' g_i$. Using the ordinary symbol of equality $=$ for elements which are equal in G_n, we have to prove that $\Pi g_i=\Pi' g_i$ implies $\Pi g_i\equiv\Pi' g_i$. The proof is made in several steps.

a. If W is any product of the g_i's, then $W\equiv g_kg_{k-1}\cdots g_1W_1$, where $0\leq k\leq n-1$ and where W_1 is a product of the generators $g_2,\cdots,g_n$ and of the elements a_i given by (5).

We prove this by induction with respect to the number m of factors $g_i^{\pm1}$ in W, since the statement is trivial in the case $m=1$ $(g_1^{-1}=g_1a_2^{-1})$. Let g $(=g_i^{\pm1})$ be the first factor of W and let W' be the product of the remaining $m-1$ factors. By our induction we can write $W\equiv gg_kg_{k-1}\cdots g_1W'_1$, where W'_1 is a product of $g_2,\cdots,g_{n-1},a_2,\cdots,a_n$. If $g=g_i^{\pm1}$, $i>k+1$, then $gg_kg_{k-1}\cdots g_1\equiv g_kg_{k-1}\cdots g_1g$ and hence $W\equiv g_kg_{k-1}\cdots g_1W_1$, where $W_1=gW'_1$. If $g=g_{k+1}$, then the product $gg_kg_{k-1}\cdots g_1W'_1$ is already of the required form. If $g=g_{k+1}^{-1}$, we have identically:

$$gg_kg_{k-1}\cdots g_1=g_{k+1}g_{k+1}^{-2}g_kg_{k-1}\cdots g_1=g_{k+1}g_k\cdots g_1a_{k+2}^{-1}$$

and hence $W\equiv g_{k+1}g_k\cdots g_1W_1$, $W_1=a_{k+2}^{-1}W'_1$. Finally, if $g=g_i^{\pm1}$, $i\leqq k$, we observe that, by (4), $(g_{i+1}g_i)^{-1}g_i(g_{i+1}g_i)\equiv g_{i+1}$, and hence

$$(g_kg_{k-1}\cdots g_1)^{-1}g_i^{\pm1}(g_kg_{k-1}\cdots g_1)\equiv(g_{i+1}g_i\cdots g_1)^{-1}g_i^{\pm1}(g_{i+1}g_i\cdots g_1)\equiv$$

$$(g_{i-1}\cdots g_1)^{-1}g_{i+1}^{\pm1}(g_{i-1}\cdots g_1)\equiv g_{i+1}^{\pm1}. \tag{7}$$

Consequently $W\equiv g_kg_{k-1}\cdots g_1g_{i+1}^{\pm1}W'_1\equiv g_kg_{k-1}\cdots g_1W_1$, q. e. d.

b. The product $g_kg_{k-1}\cdots g_1W_1$ represents a motion which carries the point P_1 into the point P_2, provided $k\geqq1$. Hence, as a corollary of *a*, it follows that *if W is a motion in which P_1 comes back to its original position, and if W is expressed as a product of the generators g_i, then $W\equiv W_1$, where W_1 is a product of the elements $g_2,\cdots,g_{n-1},a_2,\cdots,a_n$.*

c. We observe that *as a consequence of the relations* (3), (4) *and* (6) *the group generated by the elements $a_2,\cdots,a_n$ is an invariant subgroup of the group generated by the elements $g_2,\cdots,g_{n-1}$.* The proof is contained in the following relations:

$$g_ja_i\equiv a_ig_j, \qquad (j\neq1,i-1,i), \tag{8}$$

trivial if $j>i$, and immediate consequences of the relations (7), if $j\leqq i-2$;

$$\begin{aligned} g_ia_ig_i^{-1}&\equiv(g_{i-2}\cdots g_1)^{-1}g_ig^2{}_{i-1}g_i^{-1}(g_{i-2}\cdots g_1)\\ &\equiv(g_{i-2}\cdots g_1)^{-1}g_{i-1}^{-1}g_i{}^2g_{i-1}(g_{i-2}\cdots g_1)\equiv a_{i+1}; \end{aligned} \tag{8'}$$

$$\begin{aligned} g_i^{-1}a_ig_1&\equiv(g_{i-2}\cdots g_1)^{-1}g_i^{-1}g^2{}_{i-1}g_i(g_{i-2}\cdots g_1)\\ &\equiv(g_{i-2}\cdots g_1)^{-1}g_{i-1}g_i{}^2g_{i-1}^{-1}(g_{i-2}\cdots g_1)\equiv a_ia_{i+1}a_i^{-1}. \end{aligned} \tag{8''}$$

From (8′) and (8″) it follows that also $g_i^{\pm1}a_{i+1}g_i^{\mp1}$ can be expressed in terms of a_i and a_{i+1}, and hence the proof is complete.

COROLLARY. *If a product W of the g_i's represents a motion in which the point P_1 comes back to its original position, and hence in particular if $W=1$ in G_n, then $W \equiv W_g W_a$, where W_g is a product involving only the generators $g_2, \cdots, g_{n-1}$ and W_a is a product involving only the elements $a_2, \cdots, a_n$.*

d. We are now in position to prove the completeness of the relations (3), (4) and (6). We use an induction with respect to n, since in the case $n=1$ the group G_n contains only the element 1. Let W be an element of G_n, expressed in terms of the generators $g_1, \cdots, g_{n-1}$ and let $W=1$ be a true relation in G_n. By *c.*, corollary, we have $W \equiv W_g W_a$ and hence $W_g W_a = 1$. Since W_g is a motion of the points P_i in which the point P_1 is fixed, while in the motion W_a all the points P_i, except P_1, are fixed, it is clear that $W_g = 1$ must be a true relation in G_{n-1}, the generators of G_{n-1} being $g_2, \cdots, g_{n-1}$. By our induction, this relation must be a consequence of the relations (3), (4) and (6) relative to the case $n-1$. Of these, the only relation which is not included among the relations for the group G_n, is the relation (6), which for G_{n-1} is as follows: $g_2 \cdots g^2_{n-1} \cdots g_2 = 1$. Since

$$g_2 \cdots g^2_{n-1} \cdots g_2 \equiv g_1^{-2} \equiv a_2^{-1},$$

it follows that by using the relations (3), (4), (6) of the group G_n it is possible to express W_g as a product of transforms of $a_2^{\pm 1}$ by elements of G_{n-1}. By *c.*, it follows then that $W=1$ *implies a relation of the type: $W \equiv W'_a$, where W'_a is a product of the elements $a_2, \cdots, a_n$.*

We now make the following remark. The elements $a_2, \cdots, a_n$ are also generators of the Poincaré group Γ of $H - P_2 - \cdots - P_n$. Given any element V of G_n, and if V carries the point P_1 into itself, then this motion V of the n points $P_1, \cdots, P_n$ defines a deformation of the loops a_i into loops a'_{j_i}, issued from the point P_1, and *the correspondence $a_i \leftrightarrow a'_{j_i}$ defines an automorphism of the free group* Γ. It is clear that $a'_{j_i} = V^{-1} a_i V$, where now the elements a_i and a'_{j_i} are considered as elements of G_n. If the motion V is deformed continuously, while the loops a_i are fixed, then the loops a'_{j_i} are deformed continuously over $H - P_2 - \cdots - P_n$, the point P_1 remaining fixed. Hence *if $V_1 = V_2$ in G_n then $V_1^{-1} a_i V_1 = V_2^{-1} a_i V_2$ in* Γ.

Let now $V_1 = W'_a$. Since $V_1 = 1$ in G_n, it follows that $W'^{-1}_a a_i W'_a = a_i$ in Γ, i. e. W'_a is commutative with each element a_i. Since W'_a is itself an element of Γ and since the elements a_i are generators of Γ, W'_a belongs to the center of Γ. Since Γ is a free group, we must have necessarily $W'_a = 1$ in Γ, i. e. W'_a, given as a product of the elements $a_2, \cdots, a_n$, is a product of transforms of $(a_2 \cdots a_n)^{\pm 1}$. The product $a_2 \cdots a_n$, expressed in terms of the

elements g_i, is nothing but the left-hand member of (6). Hence $W'_a \equiv 1$, and consequently also $W \equiv 1$, q. e. d.

5. *The generating relations of G_n and the associated singularities of the curve C_{2n-2}.* The group G_n is the Poincaré group of the maximal cuspidal curve C_{2n-2} of order $2n-2$. It is interesting to see how the individual generating relations of G_n correspond to the singularities of C_{2n-2}. The curve C_{2n-2} possesses $3(n-2)$ cusps and $2(n-2)(n-3)$ nodes. The $(n-2)(n-3)/2$ commutativity relations (3) are the typical relations at nodes, while the $n-2$ relations (4) are the typical cusp relations (see Zariski [8]). The fact that there are 4 times as many nodes as there are relations (3) and three times as many cusps as there are relations (4), must be due partly to repetitions (two or more singularities giving one and the same relation) and partly to the fact that $g_1, g_2, \cdots, g_{n-1}$ is a reduced set of generators. The curve C_{2n-2} is of order $2n-2$, and originally a set of generators of its Poincaré group would consist of $2n-2$ loops $g_1, g_2, \cdots, g_{2n-2}$, lying in a fixed line and each surrounding a point of intersection of that line with C_{2n-2}. The curve C_{2n-2} is of class n and the n tangents of C_{2n-2} in a pencil of lines supply essentially $n-1$ equalities: $g_i = g_{2n-1-i}$ $(i = 1, 2, \cdots, n-1)$. The trivial relation $g_1 g_2 \cdots g_{2n-2} = 1$ yields the relation (6).

6. *Rational cuspidal curves of even order.* To consider a cusp of C_{2n-2} as a virtual node means to consider that cusp as a limiting case in which two critical points, or two singular lines, coincide, one corresponding to a node and the other being a simple tangent. Since the rational curves of a given order and with a given number of cusps form an irreducible system, it is immaterial which cusps of C_{2n-2} are considered as virtual nodes. We may assume therefore that any one of the relations (4), say the relation $g_1 g_2 g_1 = g_2 g_1 g_2$, is the relation at one of the cusps which are considered as virtual nodes. The above relation must then be replaced by two relations, one $g_1 g_2 = g_2 g_1$, relative to the node, and the second, $g_1 = g_2$, relative to the simple tangent. The new relation $g_1 = g_2$, combined with the relations $g_1 g_3 = g_3 g_1$ and $g_2 g_3 g_2 = g_3 g_2 g_3$, yields the relation $g_2 = g_3$. Combining this relation with the relation $g_2 g_4 = g_4 g_2$ and $g_3 g_4 g_3 = g_4 g_3 g_4$, we find $g_3 = g_4$. Continuing in this manner we get $g_1 = g_2 = \cdots = g_{n-1}$, while relation (6) becomes: $g_1^{2n-2} = 1$. We may therefore state the following result:

Any curve of even order, admitting the maximal cuspidal rational curve C of the same order as a limiting case, but possessing less cusps than C (in particular, any rational curve of even order which possesses only nodes and cusps and is not a maximal cuspidal curve) has a cyclic Poincaré group.

Let us now consider one of the nodes of C_{2n-2} as virtually non-existent, and let the relation at the node be, for instance, $g_1g_3 = g_3g_1$. The node must be then replaced by two simple tangents, very near each other, each yielding the relation $g_1 = g_3$. We must assume $n \geqq 4$, because if $n = 3$, then C_{2n-2} has no nodes. If $n \geqq 5$, the new relation $g_1 = g_3$, combined with the relations $g_1g_4 = g_4g_1$ and $g_3g_4g_3 = g_4g_3g_4$ yields the relation $g_3 = g_4$, and from this equality of two consecutive generators g_i, g_{i+1} we derive, as before, that the group is cyclic, of order $2n - 2$.

If $n = 4$, we are left with the relations:

$$g_1g_2g_1 = g_2g_1g_2, \qquad g_2g_3g_2 = g_3g_2g_3, \qquad g_1 = g_3, \qquad g_1g_2g_3{}^2g_2g_3 = 1,$$

or

$$g_1g_2g_1 = g_2g_1g_2, \qquad g_1g_2g_1{}^2g_2g_1 = 1.$$

These relations define a group $\bar{G}$, generated by two elements: $u = g_1{}^2g_2$, $v = g_1g_2$, satisfying the relations $u^2 = v^3 = 1$. In the present case the curve C_{2n-2} possesses 6 cusps (lying on a conic) and 4 nodes, and the group G_4 of C_6 reduces to the group $\bar{G}$ if one of these 4 nodes is considered as virtually non-existent. However, there is no further reduction of the group of the curve, if any or all of the 4 nodes are considered as virtually non-existent. In fact, we have shown in another paper (Zariski [8]) that $\bar{G}$ is the Poincaré group of the sextic curve with 6 cusps on a conic, and this curve is obtained from the rational sextic curve C_6 by considering its 4 nodes as virtually non-existent.[2]

Hence, we have the following results:

If a curve C of even order $2n - 2$, $n \neq 4$, and of genus > 0, admits the rational maximal cuspidal curve C_{2n-2} of the same order as a limiting case, then the Poincaré group of C is cyclic of order $2n - 2$.

In the exceptional case $n = 4$, we have the non-rational sextic curves with 6 cusps on a conic, whose Poincaré group is generated by two elements u and v, satisfying the relations $u^2 = v^3 = 1$.

7. *Rational cuspidal curves of odd order.* It has been pointed out above (section 1) that a rational maximal curve C_{2n-1}, of odd order $2n - 1$, is the dual of a rational curve Γ_{n+1}, of order $n + 1$, possessing one cusp. C_{2n-1} is a

[2] If $g_1, g_2, \cdots, g_6$ is a non-reduced set of generators for the rational sextic C_6 (see section 5), such that $g_1 = g_6$, $g_2 = g_5$, $g_3 = g_4$ are relations supplied by tangent lines of C_6, the relations at the 4 nodes of C_6 are likely to be the following: $g_1g_3 = g_3g_1$, $g_1g_4 = g_4g_1$, $g_6g_3 = g_3g_6$, $g_6g_4 = g_4g_6$. In terms of the reduced set of generators g_1, g_2, g_3, we have here only one relation repeated four times, and this explains the stablity of the Poincaré group after one node has been removed.

curve $(2n-1, 3n-5)$ and possesses $2(n-2)^2$ nodes. Let Γ_{n+1} degenerate into a rational curve Γ_n, of order n, possessing only nodes, and into a tangent line t of Γ_n, the point of contact of t with Γ_n (a tacnode of the composite curve $\Gamma_n + t$) being the limit of the cusp of the irreducible curve Γ_{n+1}. The dual curve C_{2n-1} degenerates then into the maximal cuspidal curve C_{2n-2}, the dual of the curve Γ_n, and into a tangent line p of C_{2n-2}. The composite curve $C_{2n-2} + p$ possesses as many nodes as the irreducible curve C_{2n-1} (the $2(n-2)(n-3)$ nodes of C_{2n-2} and the $2n-4$ nodes at the intersections of the line p with C_{2n-2}, outside the point of contact, hence in all $2(n-2)(n-3) + 2(n-2) = 2(n-2)^2$ nodes). Corresponding to the $3n-5$ cusps of the irreducible curve C_{2n-1}, we have on the composite curve $C_{2n-2} + p$ the $3n-6$ cusps of C_{2n-2} and the tacnode at the point of contact of p with C_{2n-2}. Inasmuch as the composite curve $C_{2n-2} + p$ is a limiting case of the curve C_{2n-1}, this tacnode must be considered as a virtual cusp. This enables us to determine the Poincaré group of the irreducible curve C_{2n-1} by investigating first the Poincaré group of the composite curve $C_{2n-2} + p$.

We consider again the space $S_n(a_0, a_1, \cdots, a_n)$ of the polynomials $a_0t^n + \cdots + a_n$, and we consider in S_n the hyperplane $S_{n-1}: a_0\xi^n + \cdots + a_n = 0$, where ξ is a fixed value of t (representative space of all the n-tuples of points on the sphere H of the complex variable t which contain the fixed point ξ). It is immediately seen that S_{n-1} touches the discriminant hypersurface Δ at every common point. This is a consequence of the elementary fact, that if $f(t)$ and $\phi(t)$ are two polynomials both divisible by $t-\xi$ and if $f(t) + \lambda_0\phi(t)$ is divisible by $(t-\xi)^2$, then the discriminant of $f(t) + \lambda\phi(t)$ is divisible by $(\lambda - \lambda_0)^2$ (in geometric language: a fixed point of a linear series $g_n{}^1$ counts for two double points of the series.) It follows that a generic plane section of the composite hypersurface $\Delta + S_{n-1}$ is exactly our composite curve $C_{2n-2} + p$, and hence the Poincaré group of $C_{2n-2} + p$ coincides with the Poincaré group of the residual space of this composite hypersurface. Let G'_n be this group. The group G'_n can be interpreted as the group of motions of n points on a sphere H with one hole, the hole being at the fixed point ξ (or also, as the group of automorphism classes of a sphere with $n+1$ holes, leaving one hole fixed). From this interpretation of G'_n, it is seen that G'_n is obtained from the group G_n by the following modifications: (a) adding to the set of generators $g_1, \cdots, g_{n-1}$ of G_n another generator γ, representing a motion of one of the points P_i, say of P_1, along a loop surrounding the point ξ and not meeting the arcs $s_2, \cdots, s_{n-1}$; (b) adding to the set of relations (3), (4) the following relations:

$$\gamma g_i = g_i \gamma, \qquad (i = 2, 3, \cdots, n-1); \tag{9}$$

(10) $$(\gamma g_1)^2 = (g_1\gamma)^2;$$

and finally, replacing the relation (6) by the following:

(11) $$\gamma g_1 \cdots g_{n-2} g^2_{n-1} g_{n-2} \cdots g_1 = 1.$$

The relations (9) and (11) are obvious. The relation (10) is obtained by observing that $g_1\gamma g_1$ can be deformed into a motion in which the point P_2 turns around both point P_1 and ξ, while the remaining points P_i remain fixed. The path of this motion does not cross the loop γ, and hence γ and $g_1\gamma g_1$ are commutative.

Having proved the completeness of the generating relations (3), (4) and (6) for the group G_n, we could prove without difficulty the completeness of the generating relations (3), (4), (9), (10), (11) for the group G'_n. However, we do not insist on the proof, since we are not immediately concerned with the group G'_n: all we need to know is that the above relations effectively hold true in G'_n.

The relation (10) is a typical relation at a tacnode, in the present case at the point of contact of the line p with C_{2n-2}. We observe incidentally that the commutativity relations (9) arise from the simple intersections of the line p with C_{2n-2}. If we regard the tacnode of the composite curve $C_{2n-2} + p$ as a virtual cusp, i. e. as the limit of a cusp and of a simple tangent line, we must replace the relation (10) by the two relations: $\gamma g_1 \gamma = g_1 \gamma g_1$ and $g_1 = \gamma$, of which the first is a consequence of the second. The new relation $g_1 = \gamma$, combined with the relations $g_1 g_2 g_1 = g_2 g_1 g_2$ and $\gamma g_2 = g_2 \gamma$, yields the relation $g_1 = g_2$. In a similar manner we find $g_2 = g_3 = \cdots g_{n-1}$. Hence *the Poincaré group of C_{2n-1} is cyclic of order $2n - 1$. A fortiori,* the group remains cyclic for any curve of order $2n - 1$ admitting C_{2n-1} as a limiting case. Hence we have the following result:

The maximal cuspidal curve C_{2n-1} of odd order $2n - 1$, and any curve of order $2n - 1$ admitting C_{2n-1} as a limiting case (in particular any rational curve of odd order possessing only nodes and cusps), possesses a cyclic Poincaré group.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] E. Artin, "Theorie der Zöpfe," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 4 (1926).

[2] A. Hurwitz, "Über Riemannsche Flächen mit gegebenen Verzweigungspunkten," *Mathematische Annalen*, vol. 39 (1891).

[3] W. Magnus, "Über Automorphismen von Fundamentalgruppen berandeter Flächen," *Mathematische Annalen*, vol. 109 (1934).

[4] K. Reidemeister, "Knotentheorie," *Ergebnisse der Mathematik und ihrer Grenzgebiete, I*, 1. Berlin, Springer (1932).

[5] B. Segre, "Esistenza e dimensione di sistemi continui di curve piane algebriche con dati caratteri," *Atti Accademia nazionale Lincei*, Rendiconti VI. s., vol. 10 (1929).

[6] F. Severi, *Vorlesungen über algebraische Geometrie*, Leipzig (1921).

[7] O. Zariski, "Algebraic surfaces," *Ergebnisse der Mathematik und ihrer Grenzgebiete, III*, 5. Berlin, Springer (1935).

[8] O. Zariski, "On the problem of existence of algebraic functions of two variables possessing a given branch curve," *American Journal of Mathematics*, vol. 51 (1929).

[9] O. Zariski, "A theorem on the Poincaré group of algebraic hypersurfaces," To appear in the *Annals of Mathematics*.

ON THE PRINCIPAL JOIN OF TWO CURVES ON A SURFACE.

By M. L. MacQueen.

1. *Introduction.* Bompiani has made [1] some important contributions to the projective differential geometry of curves in ordinary space by introducing certain lines and points, called principal lines and principal points, which are associated with the intersection of two skew curves. In connection with the investigation of the invariants of intersection of two curves, several different cases present themselves. We shall confine our attention to the case in which the two curves C, $\bar{C}$ pass through a point P with distinct tangents t, $\bar{t}$ at P and also with distinct osculating planes at P, whose line of intersection is different from t and $\bar{t}$. Bompiani shows the existence through P, in the plane determined by t, $\bar{t}$, of two lines, called *principal lines*, characterized by the property that the two cones projecting C, $\bar{C}$ from any point on either line have contact of the second or higher order along their common generator through P, instead of contact of the first order as would ordinarily be the case if the center of projection were chosen elsewhere in the plane of t, $\bar{t}$. On each principal line there is a point, called a *principal point*, which has the property that if the projecting cones have their vertices at this point, the cones have contact of the third order.

On considering the case in which the two curves belong respectively to the two families of a conjugate net on a surface, Lane has deduced [2] some interesting results. Among other things, he shows that the principal lines of the parametric curves at a point of a surface are precisely the *associate conjugate tangents* at the point. Moreover, he determines the principal points of the two parametric curves at a point of a surface, and calls the line joining these points the *principal join* of the fundamental parametric curves.

In this note we propose to supplement the investigations of Lane, by presenting other geometric characterizations of the principal join of the fundamental parametric curves at a point of a surface referred to a conjugate net. In connection with our geometric constructions we introduce the neighborhoods of the third and fourth order of the plane curves of section of the surface made

[1] E. Bompiani, "Invarianti d'intersezione di due curve sghembe," *Rendiconti dei Lincei*, ser. 6, vol. 14 (1931), pp. 456-461.

[2] E. P. Lane, "Invariants of intersection of two curves on a surface," *American Journal of Mathematics*, vol. 54 (1932), pp. 699-706.

by variable planes through the tangents of the parametric curves at a point of the surface.

2. *Power series expansions and the principal join.* Let the projective homogeneous coördinates $x^{(1)}, \cdots, x^{(4)}$ of a point P_x on a surface S referred to a conjugate net N_x in ordinary space be given as analytic functions of two independent variables u, v. The osculating planes of the parametric curves C_u, C_v at the point P_x intersect in the axis of P_x with respect to the net N_x. Let P_y be the point which is the harmonic conjugate of P_x with respect to the two foci of the axis regarded as generating a congruence when the point P_x varies over the surface. Then x and y are solutions of a completely integrable system of differential equations[3] of the form

$$
\begin{aligned}
x_{uu} &= px + \alpha x_u + Ly, \\
x_{uv} &= cx + a x_u + b x_v, \\
x_{vv} &= qx + \delta x_v + Ny
\end{aligned}
\qquad (LN \neq 0). \tag{1}
$$

From the equations

$$(x_{vv})_u = (x_{uv})_v, \qquad (x_{uu})_v = (x_{uv})_u$$

we obtain

$$y_u = fx - nx_u + sx_v + Ay, \qquad y_v = gx + tx_u + nx_v + By, \tag{2}$$

where we have placed

$$
\begin{aligned}
fN &= c_v + ac + bq - c\delta - q_u, & gL &= c_u + bc + ap - c\alpha - p_v, \\
-nN &= a_v + a^2 - a\delta - q, & tL &= a_u + ab + c - \alpha_v, \\
sN &= b_v + ab + c - \delta_u, & nL &= b_u + b^2 - b\alpha - p, \\
A &= b - (\log N)_u, & B &= a - (\log L)_v.
\end{aligned} \tag{3}
$$

The ray-points, or Laplace transformed points, x_1, x_{-1} of the curves C_u, C_v respectively at the point P_x are defined by the formulas

$$x_{-1} = x_u - bx, \qquad x_1 = x_v - ax. \tag{4}$$

The following formulas give some of the invariants of the parametric conjugate net N_x:

$$
\begin{aligned}
H &= c + ab - a_u, & K &= c + ab - b_v, \\
\mathcal{H} &= sN, & \mathcal{K} &= tL, \\
8\mathfrak{B}' &= 4a - 2\delta + (\log r)_v, & 8\mathfrak{C}' &= 4b - 2\alpha - (\log r)_u, \\
\mathfrak{D} &= -2nL, & r &= N/L.
\end{aligned} \tag{5}
$$

---

[3] E. P. Lane, *Projective Differential Geometry of Curves and Surfaces*, University of Chicago Press, 1932, p. 138.

Incidentally, it is not difficult to show that the Laplace-Darboux invariants H, K, the tangential invariants $\mathcal{H}$, $\mathcal{K}$, and the invariants $\mathfrak{B}'$, $\mathfrak{C}'$ of Green are connected by the relations

$$\mathcal{H} = H + 3\mathfrak{B}'_u + \mathfrak{C}'_v, \qquad \mathcal{K} = K + \mathfrak{B}'_u + 3\mathfrak{C}'_v. \tag{6}$$

We shall employ the covariant tetrahedron whose vertices are the points x, x_{-1}, x_1, y as a local tetrahedron of reference with a unit point chosen so that a point

$$X = y_1 x + y_2 x_{-1} + y_3 x_1 + y_4 y$$

has local coördinates $y_1, \cdots, y_4$. In this coördinate system the equations of the osculating planes of the curves C_u, C_v at the point P_x are respectively $y_3 = 0$ and $y_2 = 0$. If we introduce non-homogeneous projective coördinates by the definitions

$$x = y_2/y_1, \qquad y = y_3/y_1, \qquad z = y_4/y_1, \tag{7}$$

then a power series expansion[4] for one non-homogeneous coördinate z of a point on the surface in terms of the other two coördinates x, y is given, to terms of the fourth order, by

$$z = (1/2)(Lx^2 + Ny^2) + (4/3)(L\mathfrak{C}'x^3 + N\mathfrak{B}'y^3) + c_0x^4 + 4c_1x^3y + 4c_3xy^3 + c_4y^4 + \cdots, \tag{8}$$

where

$$c_0 = (1/3)L\mathfrak{C}'[12\mathfrak{C}' + (\log \mathfrak{C}'r^{\frac{1}{2}})_u], \qquad 4c_1 = (1/6)L(H - \mathcal{H}),$$
$$c_4 = (1/3)N\mathfrak{B}'[12\mathfrak{B}' + (\log \mathfrak{B}'r^{-\frac{1}{2}})_v], \qquad 4c_3 = (1/6)N(K - \mathcal{K}).$$

The equation

$$y = \rho z \qquad (\rho \neq 0) \tag{10}$$

represents a plane passing through the u-tangent, $y = z = 0$, at the point P_x of the surface. This plane cuts the surface in a curve whose projection from the point $(0, 0, 0, 1)$ onto the tangent plane, $z = 0$, is represented by the equation obtained by eliminating z between equations (8) and (10). If this equation is solved for y as a power series in x, the result to terms of the fourth degree is

$$y = \rho Lx^2/2 + 4\rho L\mathfrak{C}'x^3/3 + \rho(L^2N\rho^2/8 + c_0)x^4 + \cdots. \tag{11}$$

Imposing on the general equation of a conic the conditions that it be satisfied by the series (11) for y identically in x as far as the term in x^3, we obtain

[4] E. P. Lane, "A canonical power series expansion for a surface," *Transactions of the American Mathematical Society*, vol. 37 (1935), p. 481.

the equation of the conics having contact of the third order with a plane section made by a variable plane through the tangent $y=z=0$, namely,

$$y-\rho Lx^2/2-8\mathfrak{C}'xy/3+hy^2=0, \tag{12}$$

where h is a parameter. The pole of the v-tangent, $x=z=0$, with respect to any one of the conics (12) has the coördinates

$$(3/8\mathfrak{C}', 0, 0). \tag{13}$$

Similarly, if we consider a plane

$$x=\sigma z \qquad (\sigma \neq 0) \tag{14}$$

through the v-tangent, $x=z=0$, we find the equation of the conics having third order contact with the curve of section of the surface to be

$$x-\sigma Ny^2/2-8\mathfrak{B}'xy/3+kx^2=0, \tag{15}$$

where k is a parameter. The pole of the u-tangent, $y=z=0$, with respect to any one of the conics (15) is found to be the point

$$(0, 3/8\mathfrak{B}', 0). \tag{16}$$

The join of the points (13) and (16) is a line whose equation is

$$8(\mathfrak{C}'x+\mathfrak{B}'y)=3, \tag{17}$$

which is precisely the principal join of the fundamental parametric curves at the point P_x. Thus the following theorem is proved.

At each point of a surface referred to a conjugate net, the principal join of the parametric curves at the point crosses each of the parametric tangents in the pole of the other with respect to any conic having contact of the third order with the curves of intersection of the surface and the planes of a pencil with the first parametric tangent as axis.

Another geometric characterization of the principal join can be described briefly in the following way. Let us project from the ray-point $(0,0,1,0)$ onto the osculating plane, $y=0$, the curves of section of the surface (8) made by the plane (10). Eliminating y between equations (8), (10) and solving the result for z as a power series in x, the equation, in the osculating plane, $y=0$, of the projection of the curve of section is found to be

$$z=Lx^2/2+4L\mathfrak{C}'x^3/3+\cdots. \tag{18}$$

Similarly, projecting onto the osculating plane, $x = 0$, the curves of section made by the plane (14), we obtain

$$(19) \qquad z = Ny^2/2 + 4N\mathfrak{B}'y^3/3 + \cdots.$$

The conics having contact of the third order with the projections (18), (19) are respectively

$$(20) \qquad \begin{aligned} z - Lx^2/2 - 8\mathfrak{C}'xz/3 + hz^2 = 0, \\ z - Ny^2/2 - 8\mathfrak{B}'yz/3 + kz^2 = 0. \end{aligned}$$

where h, k are parameters. On finding the pole of the axis $x = y = 0$, with respect to each of the conics (20), we again obtain the points (13), (16) which determine the principal join. Thus we have proved the following theorem:

At a point P_x of a surface referred to a conjugate net let the plane curves of section of the surface made by planes passing through the tangent of the u-curve (v-curve) be projected from the ray-point of this curve onto the osculating plane of this curve. The conics in the osculating plane of the u-curve (v-curve) which have four-point contact at P_x with the projected sections determine a point on the u-tangent (v-tangent) which is the pole of the axis with respect to any one of the four-point conics of the pencil. The line which crosses the parametric tangents at the point P_x in the points thus defined is the principal join of the parametric curves at the point P_x.

3. *The quadrics of Moutard for the parametric tangents.* The osculating conic of the curve of section of the surface made by the plane (10) is contained in the pencil (12), and for this conic the parameter is found to have the value

$$(21) \qquad h = (256L\mathfrak{C}'^2 - 72c_0 - 9\rho^2L^2N)/9\rho L^2.$$

With this value of h in equation (12), the equation of the quadric of Moutard for the tangent $y = z = 0$ is found, by eliminating ρ between equations (10) and (12), to be

$$(22) \quad z = (1/2)(Lx^2 + Ny^2) + (8/3)\mathfrak{C}'xz - \left(\frac{128\mathfrak{C}'^2}{9L} - \frac{4c_0}{L^2}\right) z^2.$$

Similarly, we find the equation of the quadric of Moutard for the tangent $x = z = 0$ to be

$$(23) \quad z = (1/2)(Lx^2 + Ny^2) + (8/3)\mathfrak{B}'yz - \left(\frac{128\mathfrak{B}'^2}{9N} - \frac{4c_4}{N^2}\right) z^2.$$

The pole of the osculating plane, $x = 0$, of the v-curve with respect to the quadric (22) is found to be the point (13). Moreover, the point (16) is the

pole of the osculating plane, $y = 0$, of the u-curve with respect to the quadric (23). Thus one arrives at the following conclusion:

The principal join of the parametric curves at a point of a surface referred to a conjugate net intersects the tangent of each curve in the pole of the osculating plane of the other curve with respect to the quadric of Moutard for the tangent of the first curve.

We interpolate here a few remarks concerning the intersection of the quadrics of Moutard for the parametric tangents. It is well known, as may be verified by inspecting equations (22), (23), that these quadrics are intersected by the tangent plane of the surface in the asymptotic tangents. Furthermore, the cone projecting the curve of intersection of the two quadrics of Moutard from the point P_x consists of two planes, one of which is the tangent plane $z = 0$. The other plane, which contains the conic of intersection of the two quadrics, has the equation

$$(24)\quad \mathfrak{C}'x - \mathfrak{B}'y - \left[\frac{16}{3}\left(\frac{\mathfrak{C}'^2}{L} - \frac{\mathfrak{B}'^2}{N}\right) - \frac{3}{2}\left(\frac{c_0}{L^2} - \frac{c_4}{N^2}\right)\right] z = 0.$$

The plane (24) intersects the tangent plane of the surface in the line

$$(25)\qquad \mathfrak{C}'x - \mathfrak{B}'y = 0.$$

Moreover, the line

$$(26)\qquad \mathfrak{C}'x + \mathfrak{B}'y = 0,$$

which joins the point P_x to the point of intersection of the ray and the associate ray, has been called [5] by Davis the *second canonical tangent* of the conjugate net and its associate conjugate net. We thus obtain the following result.

The plane containing the conic of intersection of the quadrics of Moutard for the two parametric tangents at a point of a surface intersects the tangent plane of the surface in the line which is the harmonic conjugate of Davis's second canonical tangent with respect to the tangents of the net.

4. *Developables of the principal join congruence.* It is of some interest to consider the developables of the congruence of principal joins. By the usual method we find *the differential equation of the curves on the surface corresponding to the developables of the congruence of principal joins to be*

$$(27)\qquad P du^2 - (Rr - S)\, du\, dv - Qr\, dv^2 = 0,$$

[5] W. M. Davis, *Contributions to the theory of conjugate nets,* Chicago doctoral dissertation (1932), p. 19.

where the functions P, Q, R, S are defined by

$$
\begin{aligned}
P &= H + 8\mathfrak{B}'_u/3 - 64\mathfrak{B}'\mathfrak{C}'/9, \\
Q &= K + 8\mathfrak{C}'_v/3 - 64\mathfrak{B}'\mathfrak{C}'/9, \\
R &= 8\mathfrak{C}'[4\mathfrak{C}'/3 + (\log \mathfrak{C}'r^{\frac{1}{2}})_u]/3 + \mathfrak{D}/2, \\
S &= 8\mathfrak{B}'[4\mathfrak{B}'/3 + (\log \mathfrak{B}'r^{-\frac{1}{2}})_v]/3 - \mathfrak{D}r/2.
\end{aligned}
\tag{28}
$$

We propose to call the curves defined by (27) *the principal join curves* of the net N_x. Calculation of the harmonic invariant of (27) and the asymptotic curves

$$L\,du^2 + N\,dv^2 = 0 \tag{29}$$

on the surface shows, with the aid of (6), that *the principal join curves form a conjugate net if, and only if,*

$$(H - K) = 4(\mathcal{H} - \mathcal{K}). \tag{30}$$

SOUTHWESTERN COLLEGE,
MEMPHIS, TENN.

DISCONTINUOUS GROUPS ASSOCIATED WITH THE CREMONA GROUPS.

By Gerald B. Huff.

Introduction. In the theory of *complete* and *regular* linear systems of plane curves there are two closely related fundamental questions that have not been answered. First,[6] there is no simple criterion for determining when a solution $x = \{x_0; x_1, x_2, \cdots, x_\rho\}$ of the Cremona equations,

$$(1)\qquad \begin{aligned} x_1 + x_2 + \cdots + x_\rho - 3x_0 &= -d - (p-1)\\ x_1^2 + x_2^2 + \cdots + x_\rho^2 - x_0^2 &= -d + (p-1), \end{aligned}$$

actually determines a linear system $\Sigma_{p,d}$. The essential difficulty would be dissipated if the structure of the arithmetic group $g_{\rho,2}$ ([1], p. 318) were known for all ρ. $g_{\rho,2}$ is generated by

$$(2)\qquad A_{123}: \quad \begin{aligned} x'_0 &= x_0 + L,\\ x'_1 &= x_1 + L,\\ x'_2 &= x_2 + L,\\ x'_3 &= x_3 + L,\\ x'_{j+3} &= x_{j+3}, \end{aligned} \qquad \begin{aligned} &L = (x_0 - x_1 - x_2 - x_3)\\ \\ &(j = 1, 2, \cdots, \rho - 3) \end{aligned}$$

and the interchanges of $x_1, x_2, \cdots, x_\rho$. $g_{\rho,2}$ leaves invariant the linear and quadratic forms

$$(3)\qquad \begin{aligned} L &\equiv x_1 + x_2 + \cdots + x_\rho - 3x_0\\ Q &\equiv x_1^2 + x_2^2 + \cdots + x_\rho^2 - x_0^2 \end{aligned}$$

but this property is not sufficient to characterize $g_{\rho,2}$.[11]

The problem considered in this paper arises from two observations made by Coble in connection with these well-known questions. For $\rho = 9$ he found that the nature of a C-characteristic was readily determined by considering the numbers $x_0, x_1, x_2, \cdots, x_\rho$ reduced, mod 3 ([6], p. 475). We will investigate the possibility of a similar criterion for larger values of ρ. This study is made more interesting by the fact that the elements of $g_{\rho,2}$ which reduce to the identity mod 3 must constitute an invariant subgroup.[1] We will denote this invariant subgroup by $I^{(3)}_{\rho,2}$. The factor group of $g_{\rho,2}$ with respect to $I^{(3)}_{\rho,2}$ is simply isomorphic to the finite group $g^{(3)}_{\rho,2}$ obtained by reducing the coefficients in all the elements of $g_{\rho,2}$ with respect to the modulus 3.

We determine the order and nature of $g^{(3)}_{\rho,2}$ and exhibit a complete set of invariants which characterizes the group. In doing so we find that the criterion used by Coble is useful only in that single case.

I. *Definitions, conventions, and preliminary formulae.* In all the work to follow we consider the numbers $x_0, x_1, \cdots, x_\rho$ reduced modulo 3. It will be sufficient then to consider *characteristics* $x = \{x_0; x_1, x_2, \cdots, x_\rho\}$ which consist of numbers of the set 0, 1, 2. In particular,

DEFINITION. *All characteristics $\{x_0, x_1, \cdots, x_\rho\}$ which have $x_0 = i_0$ and contain l twos, m ones, and n zeros, are said to be of the type*

$$i_0; lmn, \qquad l + m + n = \rho.$$

A characteristic $\{x_0; x_1, x_2, \cdots, x_\rho\}$, which satisfies

$$\begin{aligned} x_1 + x_2 + \cdots + x_\rho - 3x_0 &\equiv a_1 \\ x_1^2 + x_2^2 + \cdots + x_\rho^2 - x_0^2 &\equiv a_2, \end{aligned} \quad \text{mod } 3, \tag{4}$$

is said to be of sort (a_1, a_2). By direct substitution in the equations (4) we obtain the theorem:

(5) *All the characteristics of sort (a_1, a_2) are included in the types*

$$0;\ (3b+\beta)(3c+\gamma)q, \quad i_0;\ (3b'+\beta')(3c'+\gamma')q', \qquad (i_0 = 1, 2)$$

where $\beta, \gamma, \beta', \gamma' = 0, 1, 2$ and satisfy the congruences

$$\begin{aligned} \beta &\equiv a_1 - a_2 \equiv \beta' + 1 \\ \gamma &\equiv -(a_1 + a_2) \equiv \gamma' + 1 \end{aligned} \quad \text{mod } 3,$$

and b, c, q, b', c', q' are any non-negative integers such that

$$3b + \beta + 3c + \gamma + q = 3b' + \beta' + 3c' + \gamma' + q' = \rho.$$

Moreover, all characteristics of these types are of sort (a_1, a_2).

Later it will be necessary to know the number of characteristics of a given sort. It is clear that there are just* $\binom{\rho}{lmn}$ distinct characteristics of type

* $\binom{\rho}{lmn}$ is the usual symbol for the coefficient of $y_1^l y_2^m y_3^n$ in the expansion of $(y_1 + y_2 + y_3)^\rho$. For l, m, n non-negative and $l + m + n = \rho$,

$$\binom{\rho}{lmn} = \frac{\rho!}{l!\,m!\,n!}$$

and is the number of distinct arrangements of ρ things of which l, m, and n are alike.

i_0; lmn. Then by (5) it follows that the number of characteristics of sort (a_1, a_2) is given by

$$(6) \qquad \sum_{b,c,q} \binom{\rho}{3b+\beta\ 3c+\gamma\ q} + 2\sum \binom{\rho}{3b'+\beta'\ 3c'+\gamma'\ q'},$$

where β, γ, β', γ' satisfy the restrictions of (5) and the sums are taken over all non-negative values such that the trinomial coefficients exist. To be able to evaluate sums such as appear in (6) we make the convention:

DEFINITION. *The nine numbers $\rho_{\beta\gamma}$ are defined for $\beta, \gamma = 0, 1, 2$ by the sum*

$$\sum \binom{\rho}{3b+\beta\ 3c+\gamma\ q},$$

where the sum is taken over all non-negative values of b, c, q such that $3b+\beta+3c+\gamma+q=\rho$.

Many interesting properties of these numbers may be derived. We include here those we must use. Since the trinomial coefficients satisfy the recursion relation

$$\binom{\rho}{l\ m\ n} = \binom{\rho-1}{l-1\ m\ n} + \binom{\rho-1}{l\ m-1\ n} + \binom{\rho-1}{l\ m\ n-1},$$

we are able to conclude immediately that the numbers $\rho_{\beta\gamma}$ satisfy the relation

$$(7) \qquad \rho_{\beta\gamma} = (\rho-1)_{\beta\gamma} + (\rho-1)_{\beta-1,\gamma} + (\rho-1)_{\beta,\gamma-1}.$$

By experiment, other algebraic forms of these numbers were found. That the results, which are tabulated below, are correct may be verified by seeing that they verify the initial conditions and the recursion relations (7).

(8) *The numbers $\rho_{\beta\gamma}$ are given by the table below:*

$$\rho \equiv 0 \quad \begin{array}{ll} \rho_{12}=\rho_{21}=3^{\rho-2}, & \rho_{00}=3^{\rho-2}-(-3)^{[\rho-1]}+(-1)^{\rho+1}(-3)^{[\rho-1]} \\ \rho_{01}=\rho_{10}=3^{\rho-2}, & \rho_{11}=3^{\rho-2} \qquad\qquad\qquad\quad +(-1)^{\rho}\ (-3)^{[\rho-1]} \\ \rho_{02}=\rho_{20}=3^{\rho-2}, & \rho_{22}=3^{\rho-2}+(-3)^{[\rho-1]} \end{array}$$

$$\rho \equiv 1 \quad \begin{array}{l} \rho_{10}=\rho_{01}=\rho_{00}=3^{\rho-2}-(-3)^{[\rho-2]}+(-1)^{\rho}\ (-3)^{[\rho-2]} \\ \rho_{12}=\rho_{21}=\rho_{11}=3^{\rho-2} \qquad\qquad\qquad +(-1)^{\rho+1}(-3)^{[\rho-2]} \\ \rho_{02}=\rho_{20}=\rho_{22}=3^{\rho-2}+(-3)^{[\rho-2]} \end{array}$$

$$\rho \equiv 2 \quad \begin{array}{l} \rho_{20}=\rho_{02}=\rho_{00}=3^{\rho-2}+(-3)^{[\rho-2]}+(-1)^{\rho+1}(-3)^{[\rho-2]} \\ \rho_{10}=\rho_{01}=\rho_{11}=3^{\rho-2} \qquad\qquad\qquad +(-1)^{\rho}\ (-3)^{[\rho-2]} \\ \rho_{21}=\rho_{12}=\rho_{22}=3^{\rho-2}-(-3)^{[\rho-2]} \end{array}$$

where $[k]$ is the largest integer in $k/2$.

Hence we consider the numbers $\rho_{\beta\gamma}$ as numbers which are readily computed and state the useful result:

(9) *The number of characteristics of sort* (a_1, a_2) *is given by* $\rho_{\beta\gamma} + 2\rho_{\beta-1,\gamma-1}$ *where* β, γ *are determined by* (5).

Having determined the types of all the characteristics of a given sort, we now seek to discover how these divide into conjugate sets under $g^{(3)}_{\rho,2}$. Since $g^{(3)}_{\rho,2}$ by definition includes the permutation group $\Pi(x_1, x_2, \cdots, x_\rho)$ and since any two characteristics of the same type are obviously conjugate under this subgroup, we seek the conditions under which two characteristics of different types are conjugate under A_{123}. In the proper sense of the word, A_{123} transforms characteristics but not types. Hence we make the convention: *we will say that the type* i_0; lmn *is conjugate to the type* i'_0; $l'm'n'$ *if any characteristic of the first type is conjugate to any characteristic of the second type under* A_{123}.* In light of this convention we readily verify the statement below.

(10) *Suppose we choose* λ *twos,* μ *ones, and* ν *zeros from the* l *twos,* m *ones, and* n *zeros of* i_0; lmn

$$\lambda + \mu + \nu = 3, \quad 0 \lesseqgtr \lambda \lesseqgtr l, \quad 0 \lesseqgtr \mu \lesseqgtr m, \quad 0 \lesseqgtr \nu \lesseqgtr m,$$

and consider a characteristic of $x_0 = i_0$ *and with the values chosen at* x_1, x_2, x_3 *in any order. Then the image of this characteristic is a characteristic of type* i'_0; $l'm'n'$ *where*

$$\begin{aligned} i'_0 &= i_0 + (i_0 - 2\lambda - \mu) \\ l' &= l + \lambda' - \lambda \\ m' &= m + \mu' - \mu \\ n' &= n + \nu' - \nu \end{aligned}$$

and λ', μ', ν' *is determined by:*

$$\lambda', \mu', \nu' \quad \textit{is equal to} \quad \lambda, \mu, \nu; \;\; \mu, \nu, \lambda; \quad \textit{or} \quad \nu, \lambda, \mu$$

according as $(i_0 - 2\lambda - \mu)$ *is* 0, 1 *or* 2.

The effect of the transformation given is to subtract out the changed numbers and add in the new ones obtained. Since this is accomplished by changing i_0 and adding numbers to l, m, n we put (10) in the following form.

---

* This means, of course, that a given type may have several conjugates under A_{123}.

(11) *The type conjugate to i_0; lmn with i_0; $\lambda\mu\nu$ isolated is*

$$
\begin{aligned}
i'_0 &= i_0 + (i_0 - 2\lambda - \mu) \\
l' &= l + L \\
m' &= m + M \\
n' &= n + N
\end{aligned}
$$

where $L = \lambda' - \lambda$, $M = \mu' - \mu$, $N = \nu' - \nu$ *with* λ', μ', ν' *determined as in* (10).

In doing the necessary computation it was found convenient to tabulate the values of i_0; LMN which arise from a chosen i_0; $\lambda\mu\nu$ as follows:

(12)

i_0	i'_0	$\lambda\mu\nu \to$	L	M	N	$\lambda\mu\nu \to$	L	M	N	$\lambda\mu\nu \to$	L	M	N
0	1	2 1 0	-1	-1	2	0 2 1	2	-1	-1	1 0 2	-1	2	-1
0	2	1 2 0	-1	-1	2	0 1 2	2	-1	-1	2 0 1	-1	2	-1
1	2	1 1 1	0	0	0								
1	2	3 0 0	-3	0	3	0 0 3	0	3	-3	0 3 0	3	-3	0
1	0	1 0 2	1	1	-2	2 1 0	-2	1	1	0 2 1	1	-2	1
2	1	1 1 1	0	0	0								
2	1	0 0 3	3	0	-3	3 0 0	-3	3	0	0 3 0	0	3	-3
2	0	0 1 2	1	1	-2	2 0 1	-2	1	1	1 2 0	1	-2	1

II. *The group $g^{(3)}_{\rho,2}$ for $\rho \leq 9$.* For $\rho \lesseqgtr 8$ the problem is greatly simplified by the fact that in these cases the order of $g_{\rho,2}$ is finite. Its order and essential properties are known.[7] Indeed, we can see that here $g^{(3)}_{\rho,2}$ and $g_{\rho,2}$ are simply isomorphic as a consequence of a theorem due to Minkowski.[8,9]

(13) *An integer linear homogeneous substitution which is of finite period and which is different from the identity cannot reduce to the identity with respect to a modulus $l \geqq 3$.*

Thus $g^{(3)}_{\rho,2}$ and $g_{\rho,2}$ are of the same order and are obviously simply isomorphic.

For $\rho = 9$ a similar procedure will suffice, even though $g_{9,2}$ is of infinite order. Dr. Taylor[3] obtained a complete determination of all the elements of $g_{9,2}$ in 1932. These results were later put in a more usable form by Dr. Barber.[2] By applying these results it is not difficult to see what happens when the coefficients in the elements of $g_{9,2}$ are reduced modulo 3. We will merely state the results obtained by direct application of the properties of $g_{9,2}$.

$g^{(3)}_{9,2}$ is of order $3^8 \cdot 2 \cdot 8640 \cdot 8\,!$. It possesses an invariant abelian subgroup of order 3^8 and type (1, 1, 1, etc.). The factor group of $g^{(3)}_{9,2}$ with respect to this invariant subgroup is simply isomorphic to $g^{(3)}_{8,2} = g_{8,2}$. $I^{(3)}_{9,2}$, the invariant subgroup of $g_{9,2}$ characterized by the fact that all its elements reduce

to the identity, mod 3, is comprised of all those elements of Dr. Taylor's[3] Type I which are defined by a $\gamma, \nu, \delta_1, \delta_2, \cdots, \delta_9$ with the property

$$\gamma \equiv \nu \equiv \delta_1 \equiv \delta_2 \equiv \cdots \equiv \delta_9, \quad \text{mod } 3.$$

This could also be stated by saying that $I^{(3)}_{9,2}$ consists of those elements[3] of a_9 which are themselves cubes of some element in a_9.

III. *The group $g^{(3)}_{\rho,2}$ for $\rho > 9$.* The essential difficulty in this paper comes in determining the nature of $g^{(3)}_{\rho,2}$ in those cases where the information concerning $g_{\rho,2}$ is fragmentary; i. e. for values of ρ greater than 9. We obtain a grip on the problem by asking how the aggregates of characteristics of the same sort divide into conjugate sets under $g^{(3)}_{\rho,2}$.

In the case $\rho = 10$ there are three self-conjugate characteristics

$$\{0;\ 2, 2, \cdots, 2\}, \quad \{0;\ 1, 1, \cdots, 1\} \quad \text{and} \quad \{0;\ 0, 0, \cdots, 0\}.$$

Excluding these, there remain 195 types of characteristics. Starting with one of these, we can find all the types conjugate to it by use of table (12). Taking the new ones we can find all their conjugates and continue the process until we have a complete conjugate set. Having done this computation, we find that for $\rho = 10$ we have exactly 9 complete conjugate sets. That is, except for the three self-conjugate characteristics, all characteristics of the same sort are actually conjugate under $g^{(3)}_{\rho,2}$. We anticipate then that this is true for all $\rho > 9$. Indeed, it may be shown that when this situation exists for $\rho - 1$, it must also exist for ρ. Suppose all the characteristics of a given sort for $\rho - 1$ are actually conjugate under $g^{(3)}_{\rho-1,2}$. The characteristics obtained by adding a zero to each of these are characteristics of that same sort for ρ and are clearly conjugate under $g^{(3)}_{\rho,2}$. It remains only to show that those characteristics of that sort which contain only twos and ones are conjugate under $g^{(3)}_{\rho,2}$ to one which contains a zero. This is readily accomplished by using table (12).

(13) *The three characteristics of types $0;\ \rho\, 0\, 0$, $0;\ 0\, \rho\, 0$, and $0, 0\, 0\, \rho$ are always self-conjugate. For $\rho > 9$, the remaining characteristics divide into just nine conjugate sets determined by the nine pairs of congruences* (4).

The significance of this is that any * two characteristics of the same sort are conjugate under $g^{(3)}_{\rho,2}$. We remark that since (5) enables us to write down all the characteristics of a given sort, we have a very complete determination

* Of course, the self-conjugate characteristics are excluded and $\rho > 9$.

of the conjugate sets for $\rho > 9$. The number of characteristics in a given conjugate set is then immediately determined by (9).

We complete the investigation of $g^{(3)}_{\rho,2}$ by comparing it with the group G_ρ. G_ρ is defined to be the group of all linear substitutions with coefficients in the $GF[3]$ which leave Q, L absolutely invariant. The immediate difficulty in comparing these two groups is that $g^{(3)}_{\rho,2}$ is defined by its generators and G_ρ is defined by the invariant forms. The difficulty is resolved by showing that:

(a) for $\rho > 9$, $g^{(3)}_{\rho,2}$ is characterized as a subgroup of G_ρ by the fact that it permutes characteristics of sort $(0,1)$ evenly; and

(b) a simple set of generators exists for G_ρ.

We make use of the notion of involutions in D-conditions introduced by Coble ([4], p. 17) and studied later by Barber.[2] Using the notation of the latter article, if $d = \{d_0;\, d_1, d_2, \cdots, d_\rho\}$ is a characteristic of sort $(0,2)$, then the linear substitution

$$(14) \qquad I_\rho(d):\quad x' = (d_i x)\, d_i + x,$$

is an involutorial element of G_ρ. By direct substitution we may verify that $I_\rho(d)$ has this useful property.*

(15) *If y is a characteristic of sort $(2,1)$ such that $y_\rho = 0$, then*

$$d = \{y_0;\, y_1, y_2, \cdots, y_{\rho-1}, 1\}$$

is of sort $(0,2)$ and the image of y under $I_\rho(d)$ is $\{0;\, 0, 0, \cdots, 0, 2\}$.

Since all characteristics of sort $(0,2)$ are conjugate under $g^{(3)}_{\rho,2}$, and since A_{123} is an involution in a D-condition, all involutions $I_\rho(d)$ are conjugate and are in $g^{(3)}_{\rho,2}$ for $\rho > 9$. By using (9) it is readily seen that $I_\rho(d)$ for $d = \{1;\, 0, 0, \cdots, 0\}$ always yields an even permutation of characteristics of sort $\{0,1\}$. Hence,

(16) *An involution in a D-condition always provides an even permutation of characteristics of sort $(0,1)$.*

Since $g^{(3)}_{\rho,2}$ is generated by involutions $I_\rho(d)$, this means that every element of the group has that property. However, there are elements of G_ρ which give an odd permutation of the characteristics in question. The involution C_{12} in the non-rational D-condition $\{\sqrt{2};\, 2\sqrt{2},\, \sqrt{2},\, 0, \cdots, 0\}$ has the equations

* This statement is also valid in the usual theory. Any P-characteristic [6] defines an involution in a D-condition which sends it into the fundamental P-characteristic $\{0;\, 0, 0, \cdots, -1\}$.

$$(17) \qquad C_{12}: \begin{aligned} x'_0 &= \qquad x_1 + 2x_2 \\ x'_1 &= 2x_0 + 2x_1 + 2x_2 \\ x'_2 &= x_0 + 2x_1 + 2x_2 \\ x'_{j+2} &= \qquad\qquad x_{j+2}, \qquad (j = 1, 2, \cdots, \rho - 2). \end{aligned}$$

Since we can count the characteristics invariant under this by use of (9), and since C_{12} interchanges the remaining ones in pairs, we show that C_{12} gives an odd permutation of characteristics of sort $(0, 1)$.

(18) *The involution C_{12} always provides an odd permutation of characteristics of sort $(0, 1)$.*

By simple direct computation it is verified that G_2 is of order 8 and is generated by involutions in D-conditions and C_{12}. As a consequence of this it may be seen that $\bar{G}_\rho$, the subgroup of G_ρ generated by involutions $I_\rho(d)$, is an invariant subgroup of index 2 under G_ρ. For by (15) an element of G_ρ which has a zero in one of its P-characteristics can be reduced to an element of $G_{\rho-1}$. The number of elements that do not have zeros is restricted and these cases may be considered in detail. This leads to the result:

(19) *G_ρ is generated by the involutions in D-conditions and C_{12}. The subgroup $\bar{G}_\rho$ generated by involutions in D-conditions is an invariant subgroup of index 2 and is completely characterized by the fact that it permutes characteristics of sort $(0, 1)$ evenly.*

The order of G_ρ is obtained by using the fact that if h is the total subgroup of H leaving an object θ invariant, then the index of h under H is the number of conjugates of θ under H ([7], p. 356; [5], p. 77). It is readily verified that $G_{\rho-1}$ is simply isomorphic to the subgroup of G_ρ which leaves $\{0; 0, 0, \cdots, 0, 2\}$ unaltered. Thus if $\theta(\rho)$ is the order of $G(\rho)$ and $C(\rho)$ is the number of conjugates of $\{0; 0, 0, \cdots, 0, 2\}$ under $G(\rho)$, then

$$(20) \qquad \theta(\rho) = C(\rho)\theta(\rho - 1).$$

By (13) all characteristics of sort $(2, 1)$ are conjugate for $\rho > 9$ and it is readily verified that this is true in the early cases. Hence by (9) we have:

(21) *$C(\rho)$, the number of conjugates of $\{0; 0, 0, \cdots, 0, 2\}$ under G_ρ, is given by $C(\rho) = \rho_{10} + 2\rho_{02} - \epsilon$ where $\epsilon = 0$ if $\rho \equiv 0, 2$, mod 3, and $\epsilon = 1$ if $\rho \equiv 1$, mod 3.*

We remark that the algebraic form of $C(\rho)$ is readily obtained from table (8). $\theta(\rho)$ is completely determined by the initial value $\theta(1)=2$, the recursion relation (20), and the fact that $C(\rho)$ is always known. From (19) it follows that $\frac{1}{2}\theta(\rho)$ is the order of $\bar{G}_\rho$ for $\rho > 2$. But for $\rho > 9$, $g^{(3)}_{\rho,2}$ contains all the generators of $\bar{G}_\rho$ and must then be simply isomorphic to $\bar{G}_\rho$.

(22) *For $\rho > 9$, the order of $g^{(3)}_{\rho,2}$ is $\frac{1}{2}\theta(\rho)$, where $\theta(\rho)$ is defined as above. $g^{(3)}_{\rho,2}$ is characterized as a subgroup of G_ρ by the fact that it permutes characteristics of sort $(0,1)$ evenly.*

Conclusion. The question concerning the possibility of distinguishing between proper and improper characteristics by reduction with respect to the modulus 3 is completely settled. Since for $\rho > 9$ all characteristics of the same sort are actually conjugate under $g^{(3)}_{\rho,2}$, it is clear that the case observed by Coble is the only one in which such procedure is useful.

The nature of $g^{(3)}_{\rho,2}$ has been determined for all $\rho > 9$. An explicit formula for the order of the group is given and it is shown that $g^{(3)}_{\rho,2}$ is completely characterized by

(a) the invariance of Q, L, and

(b) the fact that it permutes characteristics of sort $(0,1)$ evenly.*

Since G_ρ is simply isomorphic to the groups studied by Coble [12] and Dickson,[5] and since $g^{(3)}_{\rho,2}$ is an invariant subgroup of index 2 under G_ρ for $\rho > 9$, the problem of determining in detail the structure of $g^{(3)}_{\rho,2}$ has been reduced to the application of known results.

These results would be immediately applicable to the study of $g_{\rho,2}$ if the nature of $I^{(3)}_{\rho,2}$ were known for all ρ. For $\rho \lesseqgtr 9$ this was accomplished, but in the later cases only fragmentary information was exhibited. A complete characterization of $I^{(3)}_{\rho,2}$ would be necessary to give this paper a general significance.

UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] A. B. Coble, "Theta-modular groups determined by point sets," *American Journal of Mathematics*, vol. 40 (1918), p. 319.

[2] S. F. Barber, "Planar Cremona transformations," *American Journal of Mathematics*, vol. 56 (1933), p. 120.

* We remark that this imposes a condition on the elements of $g_{\rho,2}$. It is not, however, easy to apply and it is not sufficient.

[3] M. E. Taylor, "A determination of the types of Cremona transformations with not more than Nine *F*-points," *American Journal of Mathematics*, vol. 54 (1932), pp. 123-128.

[4] A. B. Coble, *American Mathematical Society Colloquium Publications*, vol. 10 (1929).

[5] L. E. Dickson, *Linear Groups* (1901), Teubner.

[6] A. B. Coble, "Cremona's diophantine equations," *American Journal of Mathematics*, vol. 56 (1934), p. 475.

[7] A. B. Coble, "Point sets and allied Cremona groups, II," *Transactions of the American Mathematical Society*, vol. 17 (1916).

[8] Hermann Minkowski, *Geometrie der Zahlen*, p. 186.

[9] Hermann Minkowski, "Zur Theorie der Positiven Quadratischen Formen," *Crelles Journal*, vol. 101 (1887), p. 196.

[10] M. Seligman Kantor, *Premier Fondements pour une Théorie des Transformations Univoques*, Naples (1871), pp. 294-295.

[11] G. B. Huff, "A note on Cremona transformations," *Proceedings of the National Academy of Sciences*, vol. 20 (1934), pp. 428-430.

[12] A. B. Coble, "Collineation groups in a finite space with a linear and a quadratic invariant," *American Journal of Mathematics*, vol. 58 (1936), p. 15.

NOTE ON ASTATIC ELEMENTS.

By F. Morley and J. R. Musselman.

Let forces in one plane have points of application a_i, magnitudes μ_i, and directions τ_i. The sum of the areas or moments about any point x is 0 when

$$\Sigma \begin{vmatrix} x & \bar{x} & 1 \\ a & \bar{a} & 1 \\ \mu\tau & \mu/\tau & 0 \end{vmatrix} = 0$$

and therefore for equilibrium

(1) $$\Sigma\mu\tau = 0$$

and

(2) $$\Sigma\mu a/\tau = \Sigma\mu\bar{a}\tau.$$

If now each force be given a turn t about its point of application, then τ_i becomes $t\tau_i$, the first equation remains, but the second becomes $\Sigma\mu a/\tau = t^2\Sigma\mu\bar{a}\tau$ and if this be true for one value of t^2 other than 1, it is always true, for $\Sigma\mu a/\tau = 0$. The condition for astatic equilibrium is then

(3) $$\Sigma\mu a/\tau = 0.$$

We take the simple case of equal forces, that is let $\mu_1 = \mu_2 = \cdots = 1$. And we speak of elements instead of forces with given points of application. We have astatic elements a_i, τ_i when

$$\Sigma\tau = 0 \quad \text{and} \quad \Sigma a_i/\tau_i = 0.$$

For instance three elements are astatic when $\tau_1 : \tau_2 : \tau_3 = 1 : \omega : \omega^2$

$$(\text{where } \omega^2 + \omega + 1 = 0)$$

and $$a_1 + \omega^2 a_2 + \omega a_3 = 0.$$

If now $\Sigma a/\tau$ is not 0, but say k, we replace a_1 by b_1 where

$$b_1/\tau_1 + a_2/\tau_2 + \cdots = 0.$$

We have then

$$a_1 - b_1 = k\tau_1$$

and similarly replacing each a_i by a b_i

$$a_i - b_i = k\tau_i$$

whence
$$\Sigma a = \Sigma b.$$

It suggests itself that the vectors $a_i - b_i$ are forces in equilibrium. This is so if

$$\Sigma \begin{vmatrix} ab \\ \bar{a}\bar{b} \end{vmatrix} = 0$$

that is if

$$\Sigma \begin{vmatrix} a, & k\tau \\ \bar{a}, & \bar{k}/\tau \end{vmatrix} = 0$$

which is true since $\Sigma a/\tau = k$. Thus if we have n elements a_i, τ_i for which

$$\Sigma\tau = 0$$

and if we replace each a_i, τ_i in turn by b_i, τ_i so as to get an astatic set, then $a_i - b_i$ form a set of equal forces in equilibrium.

If we replace $\tau_1, \tau_2, \cdots$ by $1, \epsilon, \epsilon^2 \cdots$ where ϵ is a root of $\epsilon^n = 1$, then $\Sigma\tau = 0$. The expression $\Sigma a/\tau$ becomes a Lagrange resolvent, v. The theorem becomes: If we take n elements $a_1, a_2\epsilon, a_3\epsilon^2 \cdots$ and replace in turn each a_i by b_i so that the Lagrange resolvent v vanishes for b_i and the other a's, then $a_i - b_i$ are a system of equal forces in equilibrium.

THE JOHNS HOPKINS UNIVERSITY,
AND
WESTERN RESERVE UNIVERSITY.

PROPERTIES OF THE VENERONI TRANSFORMATION IN S_4.

By Gertrude K. Blanch.

In a paper which appeared in 1901, Veneroni[1] cited that in a space S_n of n dimensions, the primals V^n_{n-1} of order n, which pass through $(n+1)$ general linear spaces S_{n-2} lying in S_n form a homaloidal system. Such a transformation is referred to as the *Veneroni* transformation. Some properties for S_n are given by Veneroni and Eiesland.[2] More specific detail about the transformation in S_4 is given by J. A. Todd[3] and by Virgil Snyder.[4] The bilinear equations defining the transformation have not heretofore been published, however. They are derived in this paper and further properties are investigated with their aid. Emphasis is laid on the study of involutorial Veneroni transformations. In S_3 any Veneroni transformations can be made involutorial by a proper choice of the frame of reference—an elegant derivation is given by H. F. Baker.[5] We show in this paper that in S_4 this is no longer true; one condition among the coefficients of the equations becomes necessary for an involution. It is found that quite generally, but not always, the involutorial case can be represented as a polarity with respect to four composite quadric primals, and the fundamental elements are considerably more specialized than in the more general involutorial transformations studied by Schoute[6] and by Alderton.[7] Furthermore, *there exist Veneroni transformations in S_4 which are involutorial; but the bilinear forms cannot be represented as polarities with respect to quadric primals by any linear transformations.* Properties which have already been found by other investigators will be included for the sake of clarity and completeness, when necessary.

Notation. Let

$$(ax) \equiv a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5;$$
$$(bx) \equiv b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5.$$

The binomial $(a_ib_k - b_ia_k)$ will be denoted by (ik). Thus $(a_1b_2 - b_1a_2) \equiv (12)$. S_n: linear space of n dimensions. V_r^n: a variety of dimension r and order n.

[1] *Lombardo Rendiconti II*, vol. 34, pp. 640-654.
[2] *Rendiconti Circolo Matematico di Palermo*, vol. 54, pp. 335-365.
[3] *Proceedings of the Cambridge Philological Society*, vol. 26 (1930), pp. 323-333.
[4] *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 673-687.
[5] *Proceedings of the London Mathematical Society* (2), vol. 21 (1923), pp. 114-133.
[6] *Archives du Musée Teyler*, ser. 2, vol. 7.
[7] *California Publications in Mathematics*, no. 15, vol. 1 (1923), pp. 345-358.

A V_r^n of S_{r+1} will be called a *primal*; if $n=1$, it will be called a *prime* and denoted by S_r. The transformation involves two spaces S_4, S'_4. The system in S_4 will be referred to as the (x)-system; that in S'_4 will be called the (x')-system. The symbol $\mathfrak{z}$ means "transform into."

1. *The derivation of the bilinear equations.* The general homaloid of this $(4-4)$ transformation in S_4 is a quartic primal V_3^4 passing through five general planes of S_4, such that any two of them intersect in one point only. It is easy to show that the system $|V_3^4|$ is ∞^4. Let the five planes have the equations:

$$\pi_1:\ x_1=x_2=0;\qquad \pi_2:\ x_3=x_4=0;\qquad \pi_3:\ x_2-x_5=x_4-x_5=0.$$
$$\pi_4:\ x_1-x_5=x_3-x_5=0;\qquad \pi_5:\ (ax)=(bx)=0.$$

The first four planes intersect by twos in the vertices of the simplex of reference and the unit point. These are entirely general for S_4. Any other six points in equally general position uniquely determine a set of four other planes which can be carried into these by a non-singular linear transformation. The fifth plane being entirely general the homaloidal system defined by these five planes will be projectively equivalent to that defined by any other five equally general planes in S_4. We demand that π_5 be non-incident[8] with the other four planes, and that the ten points of intersection of the planes π_i by twos be distinct. Then $\sum_{i=1}^{5} a_i \equiv A$; $\sum_{i=1}^{5} b_i \equiv B$ cannot both be zero; for then $(1,1,1,1,1)$ would satisfy π_5, and would coincide with the intersection $\pi_3 \cdot \pi_4$ of π_3 and π_4. It is therefore no restriction to define π_5 by $(ax)=0$; $(bx)=0$; $A \neq 0$; $\sum_{i=1}^{5} b_i = 0$.

Consider the five Segre cubic primals, each determined by four of the five planes.[9] Their equations are given below:

$$\begin{aligned}
W_1:\ & [(ax)b_2-(bx)a_2](x_5-x_2)(x_3-x_4) \\
& \quad +[(ax)b_4-(bx)a_4](x_5-x_4)(x_3-x_4)+A(bx)\cdot x_3\cdot(x_5-x_4)=0 \\
W_2:\ & [(ax)b_2-(bx)a_2\cdot](x_2-x_5)(x_2-x_1) \\
& \quad +[(ax)b_4-(bx)a_4]\cdot(x_4-x_5)(x_2-x_1)+A(bx)\cdot x_1\cdot(x_5-x_2)=0 \\
W_3:\ & [(ax)b_3-(bx)a_3]\cdot x_3\cdot(x_3-x_1) \\
& \quad +[(ax)b_4-(bx)a_4]\cdot x_4(x_3-x_1)+[(ax)b_5-(bx)a_5]\cdot x_3(x_5-x_1)=0 \\
W_4:\ & [(ax)b_3-(bx)a_3]\cdot x_3\cdot(x_4-x_2) \\
& \quad +[(ax)b_4-(bx)a_4]\cdot x_4\cdot(x_4-x_2)+[(ax)b_5-(bx)a_5]\cdot x_4\cdot(x_5-x_2)= \\
W_5:\ & x_1x_3x_4-x_1x_2x_3+x_1x_2x_4-x_1x_4x_5+x_2x_3x_5-x_2x_3x_4=0.
\end{aligned}$$

---

[8] Two planes are *non-incident* in S_4 when they have only a point in common; if they have a line in common, call them *incident*.

[9] Bertini, *Einführung in die projective Geometrie mehrdimensionaler Räume*, chapter VIII, p. 193.

It will be useful to recall some of the properties of these primals. Each can be generated as the locus of lines which meet the four planes determining it. Each contains fifteen planes; from every one of its points there can be drawn six lines, each such line meeting six of the fifteen planes. We may readily verify the following identity:

$$(x_1 - x_2) W_1 - (x_4 - x_3) W_2 + (bx) A W_5 \equiv 0.$$

Let

(1.1) $\quad x'_1 = x_1 W_1;\ x'_2 = x_2 W_1;\ x'_3 = x_3 W_2;\ x'_4 = x_4 W_2;\ x'_5 = A(ax) W_5.$

Each x'_i is satisfied by all five planes π_i and is therefore a composite member of $|V_3^4|$; further these five homaloids are linearly independent and hence they define a proper Veneroni transformation. To obtain now the bilinear equations consider the determinant

$$(1.2)\quad \begin{vmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 & \lambda_5 \\ -x_2 & x_1 & 0 & 0 & 0 \\ 0 & 0 & -x_4 & x_3 & 0 \\ (ax) & -(ax) & -(ax) & (ax) & (bx) \\ \begin{matrix}[(a_2 - A)(x_5 - x_2) \\ + a_4(x_5 - x_4)]\end{matrix} & \begin{matrix}[a_2(x_2 - x_5) \\ + a_4(x_4 - x_5)]\end{matrix} & \begin{matrix}[a_2(x_2 - x_5) \\ + (A - a_4)(x_5 - x_4)]\end{matrix} & \begin{matrix}[a_2(x_5 - x_2) \\ + a_4(x_5 - x_4)]\end{matrix} & \begin{matrix}[b_2(x_5 - x_2) \\ + b_4(x_5 - x_4)]\end{matrix} \end{vmatrix} = 0.$$

This expands into

$$\lambda_1 x_1 W_1 + \lambda_2 x_2 W_1 + \lambda_3 x_3 W_2 + \lambda_4 x_4 W_2 + \lambda_5 A(ax) W_5 = 0.$$

The determinant is therefore the complete system $|V_3^4|$. Replace the first row of the determinant successively by the second, third, fourth, and fifth rows; taking account of (1.1), we have the four bilinear equations:

M_1: $x_1 x'_2 - x_2 x'_1 = 0$

M_2: $x_3 x'_4 - x_4 x'_3 = 0$

M_3: $(ax)(x'_1 - x'_2 - x'_3 + x'_4) + (bx) x'_5 = 0$

M_4: $(x_5 - x_2) \cdot (a_2 C - A x'_1 + b_2 x'_5) + (x_5 - x_4) \cdot (a_4 C + A x'_3 + b_4 x'_5) = 0.$

The linear combination $A(M_1 - M_2) + M_3 + M_4$ gives

$$M_5:\ (x_5 - x_1) \cdot (-a_1 C - A x'_2 - b_1 x'_5) + (x_5 - x_3) \cdot (-a_3 C + A x'_4 - b_3 x'_5) = 0$$

where $C \equiv x'_1 - x'_2 - x'_3 + x'_4$.

The matrix of each bilinear equation has rank 2; the direct and inverse transformation are of the same nature; and the fundamental planes of (x) and (x') can be read off from inspection of these equations. Naming the fundamental planes of (x') T_1, T_2, T_3, T_4, T_5, we may say that each bilinear

equation associates a plane T_i with a plane π_i; and we shall hereafter call two such planes T_i, π_i satisfying the same bilinear equation *associated* planes.

3. *Locus of invariant points.* When the spaces (x) and (x') are superposed, the invariant points satisfy M_1 and M_2 identically. The locus of invariant points therefore consists of the quartic surface defined by M_3 and M_4, with the primed symbols removed. The planes π_3, π_4, π_5 intersect the surface in conics, but π_1, π_2 cut the surface in only four points. Thus the invariant locus is not symmetric with respect to the fundamental planes and depends on the choice of the frame of reference.

Involutorial Transformations.

4. *A special involutorial transformation.* The transformation already developed will be involutorial when and only when $(ax) \equiv x_5$ and $(bx) \equiv -C$. The fundamental system becomes very much specialized—$R_2{}^5$ breaks up into two planes and an $R_2{}^3$, and four of the Segre cubic primals on $J_3{}^{15}$ contain double lines, with only 11 planes on each. The transformation

$$x'_1 = y_2;\ x'_2 = y_1;\ x'_3 = y_4;\ x'_4 = y_3;\ x'_5 = y_5 \tag{2.1}$$

applied to (x') changes these forms to a polarity with respect to four composite quadric primals.

5. *Veneroni transformations in the form of quadric polarities.* Subjecting (x') to a linear transformation is equivalent to taking for the simplex of reference in (x') different linear combinations of the homaloids in (x). When the two spaces are regarded as superposed, $J_2{}^{10}$, $J'_2{}^{10}$ have two planes in common, each of which is a W_i. Hence the square of the derived transformation is of order ten. But we may transform any four planes of (x') into any other four planes in equally general position and we may so choose such a transformation that four, three, two, one or no planes T_i of (x') coincide with fundamental planes of (x). Hence the square of the general Veneroni transformation may be of order 4, 7, 10, or 16 respectively, depending merely on the frame of reference. But under what minimum restrictions will the transformation be involutorial? We have studied one involutorial transformation in which the fundamental system was very much specialized. We will now obtain others much more general. In any involutorial transformation, the fundamental planes of (x) and (x') must coincide; but conceivably they may be associated with one another in any one of the 5! ways in the bilinear equations. If the plane T_i coincides with its *associated* plane π_i for every i,

we shall call such a transformation *form I*. We inquire under what conditions form I is possible, with at least four planes T in perfectly general position. In that case there will exist a non-singular linear transformation on (x') carrying the four planes T_i into π_i, which will also transform T_5 into π_5. Then such a transformation must have the form:

$$x'_1 = \alpha_1 y_1 + \alpha_2 y_2;\quad x'_2 = \alpha_3 y_1 + \alpha_4 y_2;\quad x'_3 = \alpha_5 y_3 + \alpha_6 y_4;\quad x'_4 = \alpha_7 y_3 + \alpha_8 y_4;$$
$$x'_5 = \alpha_9 (ay) + \alpha_{10} (by).$$

In addition, it must carry the remaining planes T_i into the associated planes. Completing the algebraic details[10] we find that *one restriction* on the coefficients is necessary; namely

$$(2.2)\qquad (24)\cdot(51)b_3 + (25)\cdot(13)b_4 = 0, \quad \text{when} \quad \sum_{i=1}^{5} b_i = 0.$$

If, in addition

$$(2.3)\quad [(53)\cdot(51)\cdot(42)\cdot(31)b_2 b_4]\cdot[(24)b_1 + (13)b_2 + Ab_1 b_2]\cdot[(13)b_4 + (42)b_3 - Ab_3 b_4] \neq 0$$

the transformation will always be non-singular. If some of the factors in (2.3) are zero, form I may still be possible in some cases, and impossible in others. The special restrictions in individual cases can be derived quite readily. The resulting bilinear equations will show *that when form I is possible, the transformation is always a polarity with respect to four composite quadrics* (and indeed, a fifth also, when M_5 is considered). No essential restriction is imposed on the fundamental elements. Thus the double points are distinct and no three need be collinear; the fifty planes of $J_3{}^{15}$ can all be distinct. In S_3, every Veneroni transformation can be made involutorial by a suitable linear transformation; and since so much depends on the frame of reference, we will examine more closely the choice used to see whether (2.2) is really an essential restriction for S_4. First, we imposed the restriction that $\Sigma b_i = 0$ in developing the bilinear equations. It may be verified[11] that when this restriction is removed, condition (2.2) takes the form

$$B[(51)\cdot(24)a_3 + a_4(31)\cdot(52)] + A[b_4(13)\cdot(52) + b_3(51)\cdot(42)] = 0,$$

where $B \equiv \Sigma b_i$. This is a relative invariant under the transformations $(ay) = c(ax) + d(bx)$; $(by) = e(ax) + g(bx)$. It remains to be seen whether by imposing a transformation (ω) on (x) and another transformation

[10] The algebra in full detail is given in the author's dissertation, 1935, Library of Cornell University.

[11] Author's dissertation, Appendix I b.

(τ) on (x'), we may obtain form I without any restriction on the coefficients a_i, b_i. Suppose one such had been found. Then the fundamental planes of (x) and (x') now being alike, use (ω^{-1}) on both. Then $(\tau\omega^{-1})$ transforms all five planes of (x') into the associated planes of (x) under the old frame of reference, and if form I is possible at all, we may obtain it by keeping (x) fixed and impositing a linear transformation on (x'). *Hence condition (2.2) is a necessary restriction.*

6. *Involutorial Veneroni transformations which cannot assume form I.* The first special involutorial transformation in II is not a polarity with respect to quadric primals, since $a_{ik} \neq a_{ki}$ in the fourth bilinear equation. But condition (2.2) is satisfied and (2.1) carries the transformation into form I. Is it possible to find involutorial Veneroni transformations, not form I, and such that form I is never possible for them? Such transformations exist and we give a method of obtaining them by examining certain anharmonic ratios.

From every double-point $\pi_i\pi_k$ can be drawn a line of R_2^5 to meet the other three fundamental planes. Let A_3, A_4, A_5 be the points in π_3, π_4, π_5 respectively on the line of R_2^5 from $\pi_i\pi_2$. Let $p_{12} = (\pi_i\pi_2, 3, 4, 5)$ be the anharmonic ratio of the points $\pi_i\pi_2, A_3, A_4, A_5$. Obtain the anharmonic ratios p_{ik} on every one of the ten lines through the points $\pi_i\pi_k$. Similarly, obtain the corresponding cross-ratios p'_{ik} on the lines in (x') from the associated points T_iT_k, in the same order. *Then it is readily verified that $p_{ik} = p'_{ik}$ for every i, k.*

Now suppose the five planes T_i are projective with the planes π_j but not in the associated order. Suppose, for instance, there exists a linear transformation under which $T_1 \,\mathcal{Z}\, \pi_1$; $T_2 \,\mathcal{Z}\, \pi_2$; $T_5 \,\mathcal{Z}\, \pi_5$; $T_3 \,\mathcal{Z}\, \pi_4$: $T_4 \,\mathcal{Z}\, \pi_3$. Since cross-ratios are invariant under a linear transformation, we must have $(\pi_1\pi_2, 3, 4, 5) = (\pi_1\pi_2, 4, 3, 5)$. Hence $p_{12} = 1/2$. The other anharmonic ratios impose three more independent conditions, and we exhibit the following equations for π_5 which will satisfy all requirements:

$$(ax) : (2b+1)x_1 + b(2b+1)x_2 + 2(b^2+b+1)x_3 - 2(b^2+b+1)x_5 = 0$$

$$(bx) : (-2b-1)x_1 + (2b+1)x_2 + 2bx_3 - 2(b+1)x_4 + 2x_5 = 0$$

subject to the condition $(2b+1)(b+1)(b^2+b+1) \neq 0$, but b unrestricted otherwise. The following linear transformation carries $T_i \,\mathcal{Z}\, \pi_j$:

$$x'_1 = a_1^2(y_1 + by_2)/a_3;\ x'_2 = a_1^2(by_1 + y_1 - y_2)/a_3;\ x'_3 = y_3 + 2by_4$$
$$x'_4 = 2(b+1)y_3 - y_4;\ x'_5 = (a_1/2)[a_1y_1 + 2(y_3 - y_5) + 2b(y_4 - y_5)].$$

where $a_1 = 2b + 1$; $a_3 = 2(b^2 + b + 1)$.

The five bilinear equations become:

$$M_1:\ x_1[(b+1)y_1-y_2]-x_2(y_1+by_2)=0$$
$$M_2:\ x_3[(2b+2)y_3-y_4]-x_4(y_3+2by_4)=0$$
$$M_3:\ (ax)[(ay)+(b^2+b+1)\cdot(by)] + (bx)[(ay)-b(by)]\cdot(b^2+b+1)=0$$
$$M_4:\ (x_5-x_2)[(2b+1)\cdot(y_3-y_5)] + (x_5-x_4)[(2b+1)\cdot(y_5-y_1)+(y_5-y_3)]=0$$
$$M_5:\ (x_5-x_1)\cdot[(2b+1)(y_5-y_4)] + (x_5-x_3)\cdot[(2b+1)(y_2-y_5)+(y_5-y_4)]=0.$$

The transformation is non-singular when $(b^2+b+1)(b+1)(2b+1)\neq 0$. It is readily verified that

$$(24)(51)b_3+(25)(13)b_4=-8(2b+1)^3\cdot(b+1)^2(b^2+b+1)$$

and is never zero when the transformation is non-singular. Hence form I is never possible for these transformations.

In existing literature, involutorial transformations in S_4 expressible by means of bilinear equations are given as polarities with respect to quadric primals by Schoute [12] and by Alderton.[13] The results of this paper show that not all involutorial transformations so definable are expressible as quadric polarities. These Veneroni transformations also throw more light on the character of the fundamental system for the case of the quadric polarities. Miss Alderton classifies the transformations according to the characteristics of the quadric primals defining the polarity. For the general case of all four primals being "space-pairs" (composite primals), she cites that J_2^{10} consists of four planes and a sextic surface, with more specialization when the planes common to the space-pairs are incident. The Veneroni transformations show that J_2^{10} can be considerably more specialized even when the four "space-pairs" are perfectly general.

Clearly a transformation may be involutorial even when $a_{ik}\neq a_{ki}$ in all the bilinear equations. A less exacting criterion—quite obvious yet worth noting is this: Denote by the (x') matrix the one from which the ratios $x'_1:x'_2:x'_3:x'_4:x'_5$ are obtained. If the (x) matrix may be obtained from the (x') matrix by elementary transformations on the rows of the latter, then the transformation is involutorial. The cases given in this paper satisfy this criterion.

CORNELL UNIVERSITY.

[12] Schoute, *loc. cit.*

[13] Alderton, *loc. cit.*

SECOND ORDER DIFFERENTIAL EQUATIONS WITH TWO POINT BOUNDARY CONDITIONS IN GENERAL ANALYSIS.[1]

By A. D. MICHAL and D. H. HYERS.

Introduction. The abstract differential calculus initiated by M. Fréchet[2] has led several authors to the study of both "ordinary"[3] and "total"[4] differential equations in abstract normed vector spaces. So far the existence theorems given have all been for one point initial conditions.

In this paper we consider the following "ordinary" second order differential equation with *two point* boundary conditions:

$$(1) \qquad d^2\xi(t)/dt^2 = F(t, \xi, d\xi/dt), \qquad \xi(t_0) = A, \quad \xi(t_1) = B,$$

where the values of the functions $\xi(t)$, $F(t, \xi, \eta)$ and the second and third arguments of $F(t, \xi, \eta)$ are elements of a Banach space[5] while t is a real variable.

Section 1 contains essentially an abstraction of Picard's two point existence theorem.[6] In section 2 we consider special systems of type (1) (Cf. systems (2.1) and (2.6)) occurring in abstract differential geometry,[7] and study the continuity and Fréchet differentiability of the solution $x(t, x_0, x_1)$ as a function of t and the boundary values x_0, x_1. The main results of the

---

[1] Presented to the American Mathematical Society, September, 1935, and November, 1935.

[2] M. Fréchet, *Comptes Rendus*, vol. 180 (1925), pp. 806-809; *Annales Scientifique de l'École Normale Supérieure*, vol. 42 (1925), pp. 293-323. For the differentials of functionals see V. Volterra, *Theory of Functionals* (London, 1930); G. C. Evans, *American Mathematical Society Colloquium Lectures* (New York, 1918).

[3] L. M. Graves, *Transactions of the American Mathematical Society*, vol. 29 (1927), p. 514; M. Kerner, *Prace Matematyczno-Fizyczne*, vol. 40 (1933), pp. 47-67. For the earlier work on linear differential equations in Moore's general analysis see *E. H. Moore, Atti di IV Congresso* (Rome, 1908), vol. 2, pp. 98-114; T. H. Hildebrandt, *Transactions of the American Mathematical Society*, vol. 18 (1917), p. 73.

[4] A. D. Michal and V. Elconin, *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 534-536; A. D. Michal, *Annali di Matematica* (in press).

[5] S. Banach, *Opérations Linéaires* (Warsaw, 1932). In (1) we have written $\dfrac{d^2\xi(t)}{dt^2}$ in the place of $\dfrac{d}{dt}\left(\dfrac{d\xi(t)}{dt}\right)$. For other definitions of second derivatives see M. Kerner, *Annali di Matematica*, vol. 10 (1932), pp. 145-164.

[6] E. Picard, *Traité d'Analyse*, Tome 3 (Paris, 1928), pp. 90-96.

[7] A. D. Michal, *Annali di Matematica, loc. cit.*; *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 526-529.

paper center around Theorems 2, 3, 4 and 5 of this section. Sections 3 and 4 are concerned with a differential equation in a normed ring and with an integro-differential system arising in functional geometry.[8]

1. *A general existence theorem.* In this section we give an existence theorem for the differential system (1) together with a lemma needed in the proof of this theorem and of Theorem 6.

LEMMA 1. *Let $\Phi(t)$ be a continuous function on the closed real interval $(0, c)$ to a Banach space and let*

$$m = \max_t \| \Phi(t) \|.$$

Then for $0 \leq t \leq c$ the unique solution of the differential equation

$$d^2\xi/dt^2 = \Phi(t) \tag{1.1}$$

taking on the two-point conditions $\xi(0) = 0$, $\xi(c) = B$ is given by

$$\xi(t) = \int_0^c g(s, t)\Phi(s)\,ds + (t/c)B,$$

where

$$\begin{aligned} g(s, t) &= s(t/c - 1) \qquad s < t \\ &= t(s/c - 1) \qquad s > t. \end{aligned}$$

Furthermore we have

$$\| \xi \| < mc^2/8 + \| B \|, \qquad \| d\xi/dt - B/c \| < mc/2.$$

THEOREM 1. *Let L and L' be any two chosen positive numbers. Denote the real closed interval (a, b) by $\boldsymbol{I}$, let $\boldsymbol{N}_1$ be the set of points ξ of a Banach space*[9] *$\boldsymbol{E}$ such that*

$$\| \xi - A \| \leq L$$

and let N_2 be the set of points η of E such that

$$\| \eta - \eta_0 \| \leq L'.$$

Suppose $F(t, \xi, \eta)$ to be a continuous function in the set t, ξ, η for t in $\boldsymbol{I}$, ξ in $\boldsymbol{N}_1$, η in $\boldsymbol{N}_2$, and with values in $\boldsymbol{E}$. Further suppose that for the same set of values of t, ξ, η, the function F satisfies the Lipschitz condition

[8] A. D. Michal, *American Journal of Mathematics*, vol. 50 (1928), pp. 473-517; *Proceedings of the National Academy of Sciences*, vol. 16 (1930), (three papers) and vol. 17 (1931).

[9] Banach spaces and their subsets will be denoted throughout the paper by bold face letters.

$$\| F(t, \xi_1, \eta_1) - F(t, \xi_2, \eta_2) \| \leq \alpha \| \xi_1 - \xi_2 \| + \beta \| \eta_1 - \eta_2 \|$$

(and hence $\| F \|$ *is bounded, say* $\| F(t, \xi, \eta) \| \leq U$*). Then subject to the following inequalities*

$$(1.2) \qquad \begin{cases} a \leq t_0 < t_1 \leq b \\ \dfrac{U(t_1 - t_0)^2}{8} + \| B - A \| < L \\ \dfrac{U(t_1 - t_0)}{2} + \left\| \eta_0 - \dfrac{B - A}{t_1 - t_0} \right\| < L' \\ \dfrac{\alpha(t_1 - t_0)^2}{8} + \dfrac{\beta(t_1 - t_0)}{2} < 1, \end{cases}$$

the differential system (1) *has a solution on* (t_0, t_1). *Furthermore under the same conditions there exists on* (t_0, t_1) *a unique solution* $\xi(t)$ *of the differential system* (1) *and the inequalities*

$$(1.3) \qquad \| \xi(t) - A \| \leq L, \qquad \| d\xi(t)/dt - \eta_0 \| \leq L'.$$

The proof of Lemma 1 and the existence of a solution of the differential system (1) can be obtained by an evident modification of Picard's[10] well known methods and by the application of some known or easily proved results on abstractly valued functions.[11] To prove the uniqueness of the solution of (1) and (1.3) we take $t_0 = 0$ without loss of generality. Assume there are two distinct solutions $\xi(t)$ and $\eta(t)$ in $(0, t_1)$. By the Lipschitz condition we obtain the inequality

$$\| F(t, \xi(t), d\xi(t)/dt) - F(t, \eta(t), d\eta(t)/dt) \| \\ \leq \alpha \| \xi(t) - \eta(t) \| + \beta \| d\xi(t)/dt - d\eta(t)/dt \|$$

for t in $(0, t_1)$. The non-negative continuous function on the right side of this inequality takes on its positive maximum value, say N, at some point τ in $(0, t_1)$. Hence an application of the lemma to the differential equation

$$\frac{d^2(\xi(t) - \eta(t))}{dt^2} = F\left(t, \xi(t), \frac{d\xi(t)}{dt}\right) - F\left(t, \eta(t), \frac{d\eta(t)}{dt}\right)$$

yields

$$N < N(\alpha t_1^2/8 + \beta t_1/2).$$

But this contradicts the last inequality (1.2) for $t_0 = 0$.

2. *A differential equation arising in abstract differential geometry.* One

[10] E. Picard, *loc. cit.*

[11] M. Kerner, *Prace Matematyczno-Fizyczne, loc. cit.*

of us[12] has initiated the study of Riemannian and non-Riemannian differential geometries in abstract vector spaces. In these geometries the geodesics (or paths) satisfy special differential equations of type (1). In this section we shall prove several theorems on geodesics.

LEMMA 2. *Let* $\Gamma(x, \xi, \eta)$, *an affine connection,*[12] *be defined on*[13] $\boldsymbol{E}((\bar{x})_a)\boldsymbol{E}^2$ *to a Banach space* $\boldsymbol{E}$. *If*

(i) $\Gamma(x, \xi, \eta)$ *is additive in* ξ *and in* η;

(ii) $\| \Gamma(x, \xi, \eta) \| \leq M \| \xi \| \ \| \eta \|$ *(M, constant)*;

(iii) *The partial Fréchet differential*[14] $\Gamma(x, \xi, \eta; \delta x)$ *exists and is continuous in* x, *then there exist positive numbers* N *and* $\bar{a} < a$ *such that for any* λ

$$\| \Gamma(x_1, \xi_1, \xi_1) - \Gamma(x_2, \xi_2, \xi_2) \| \leq N\lambda^2 \| x_1 - x_2 \| + 2M\lambda \| \xi_1 - \xi_2 \|$$

for all x *in* $(\bar{x})_{\bar{a}}$ *and all* ξ *in* $(0)_\lambda$.

Proof. Clearly $\Gamma(x, \xi, \eta)$ is bilinear[15] in ξ and η. On using hypothesis (iii) and a theorem on differentials proved by one of us[16] we find that $\Gamma(x, \xi, \eta; z)$ is trilinear in ξ, η, z.

Therefore there exists[17] an $\bar{a}$, $0 < \bar{a} < a$ and a constant N such that

$$\| \Gamma(x, \xi, \eta; z) \| \leq N \| \xi \| \ \| \eta \| \ \| z \|$$

for x in $(\bar{x})_{\bar{a}}$. For any positive λ let x_1 and x_2 be in $(\bar{x})_{\bar{a}}$ and ξ_1, ξ_2 be in $(0)_\lambda$. On using the formula[18] for the difference in terms of the differential and noting that a neighborhood is a convex point set we obtain

$$\| \Gamma(x_1, \xi_1, \xi_1) - \Gamma(x_2, \xi_1, \xi_1) \| \leq N\lambda^2 \| x_1 - x_2 \|.$$

---

[12] A. D. Michal, *Annali di Matematica, loc. cit.*; *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 526-529.

[13] $\boldsymbol{E}((\bar{x})_a)$ or simply $(\bar{x})_a$ is the set of points of $\boldsymbol{E}$ for which $\| x - \bar{x} \| < a$. By "$f(x_1, \cdots, x_n)$ on $\boldsymbol{E}_1\boldsymbol{E}_2 \cdots \boldsymbol{E}_n$ to $\boldsymbol{E}$" we mean that the i-th argument x_i ranges over the set $\boldsymbol{E}_i$ while the values of $f(x_1, \cdots, x_n)$ are in $\boldsymbol{E}$.

[14] M. Fréchet, *loc. cit.* See also L. M. Graves and T. H. Hildebrandt, *Transactions of the American Mathematical Society*, vol. 29 (1927).

[15] Additive and continuous in ξ and η separately.

[16] A. D. Michal, *Annali di Matematica, loc. cit.*

[17] M. Kerner, *Annals of Mathematics*, vol. 34 (1933), pp. 546-572.

[18] L. M. Graves, *Transactions of the American Mathematical Society*, vol. 29 (1927), p. 173. See also M. Kerner, *Prace Matematyczno-Fizyczne*, and *Annals of Mathematics*, *loc. cit.*

From hypotheses (i) and (ii) we have

$$\| \Gamma(x_2, \xi_1, \xi_1) - \Gamma(x_2, \xi_1, \xi_2) \| \leq M\lambda \| \xi_1 - \xi_2 \|$$
$$\| \Gamma(x_2, \xi_1, \xi_2) - \Gamma(x_2, \xi_2, \xi_2) \| \leq M\lambda \| \xi_1 - \xi_2 \|$$

and the lemma follows immediately.

So far we have considered two point boundary problems in which one of the points x_0 is kept fixed while the other point x_1 is any chosen point within a suitable domain. We shall now give some theorems on geodesics in which both end points x_0 and x_1 are any chosen points within a suitable domain.

THEOREM 2. *Suppose the hypotheses of Lemma 2 are satisfied. Using the notations of this lemma choose* $\lambda <$ *the smaller of*

$$\frac{\bar{a}}{2(t_1 - t_0)}, \quad \frac{1}{M(t_1 - t_0)}, \quad \frac{2}{N(t_1 - t_0)} \left(\sqrt{4M^2 + 2N} - 2M\right)$$

and choose $\delta < (\lambda/4)(t_1 - t_0)$. *Then for any such choice of* λ *and* δ, *and for any* x_0, x_1 *in* $(\bar{x})_\delta$, *the system*

$$(2.1) \qquad d^2x/dt^2 = \Gamma(x, dx/dt, dx/dt), \qquad x(t_0) = x_0, \quad x(t_1) = x_1$$

has a solution which with its derivative is uniformly continuous in t, x_0, x_1. *Furthermore there exists on* (t_0, t_1) *a unique solution* $x(t)$ *of the system consisting of* (2.1) *and*

$$\| x(t) - \bar{x} \| \leq \bar{a}, \qquad \| dx/dt \| \leq \lambda.$$

Proof. From Lemma 2 and the inequality $2\delta < \bar{a}/4$ we see that Theorem 1 applies here if we take $L = \bar{a}/2$, $L' = \lambda$, $\eta_0 = 0$, $\alpha = N\lambda^2$, $\beta = 2M\lambda$, $U = M\lambda^2$. From the hypotheses on λ it follows that the inequalities (1.2) of Theorem 1 are satisfied. A glance at the uniqueness proof of Theorem 1 shows the truth of the last statement in our theorem. To prove the uniform continuity write $t_0 = 0$, $t_1 = c$ and let

$$(2.2) \qquad \begin{cases} y_0(t, x_0, x_1) = x_0 + (t/c)(x_1 - x_0) \\ y'_0(t, x_0, x_1) = (1/c)(x_1 - x_0). \end{cases}$$

Consider the following system of integral equations which is easily shown to be equivalent to differential system (2.1):

$$(2.21) \qquad \begin{cases} x(t, x_0, x_1) = y_0(t, x_0, x_1) + \int_0^c g(s, t)\Gamma(x(s), x'(s), x'(s))\,ds \\ x'(t, x_0, x_1) = y'_0(t, x_0, x_1) + \int_0^c g_t(s, t)\Gamma(x(s), x'(s), x'(s))\,ds. \end{cases}$$

The functions y_i and y'_i given recurrently below approximate to the solution $x(t, x_0, x_1)$ and its derivative respectively:

$$(2.3)\qquad \begin{cases} y_i\,(t, x_0, x_1) = y_0\,(t, x_0, x_1) + \displaystyle\int_0^c g(s, t)\Phi_{i-1}(s, x_0, x_1)\,ds \\ y'_i(t, x_0, x_1) = y'_0(t, x_0, x_1) + \displaystyle\int_0^c g_t(s, t)\Phi_{i-1}(s, x_0, x_1)\,ds \\ \qquad\qquad (i = 1, 2, \cdots) \end{cases}$$

where

$$\Phi_i(s, x_0, x_1) = \Gamma(y_i(s, x_0, x_1), y'_i(s, x_0, x_1), y'_i(s, x_0, x_1)).$$

By means of the inequailty

$$(2.31)\qquad \begin{cases} \| \Phi_i(s, x_0 + \delta x_0, x_1 + \delta x_1) - \Phi_i(s, x_0, x_1) \| \\ \qquad \leq N\lambda^2 \| y_i(s, x_0 + \delta x_0, x_1 + \delta x_1) - y_i(s, x_0, x_1) \| \\ \qquad + 2M\lambda \| y'_i(s, x_0 + \delta x_0, x_1 + \delta x_1) - y'_i(s, x_0, x_1) \| \end{cases}$$

and an induction we see that y_i and y'_i are uniformly continuous in t, x_0, x_1 for t in $(0, c)$ and x_0, x_1 in $(\bar{x})_\delta$. The uniform continuity of the limit functions follows by the usual method.

To investigate the differentiability of the solution $x(t, x_0, x_1)$ as a function of the set of boundary values (x_0, x_1) we could proceed directly as in the proof of continuity of Theorem 2. An alternative method,[19] which we employ below, is to make use of the general implicit function theorems of Hildebrandt and Graves.

Before studying the differentiability of $x(t, x_0, x_1)$ we prove the following lemma concerning functions[20] of class $C^{(n)}$.

LEMMA 3. *Let* $\mathbf{E}$ *and* $\mathbf{E}_1$ *be Banach spaces and* $\mathbf{X}$ *be a bounded convex region of* $\mathbf{E}$. *The necessary and sufficient conditions that a function* $F(x)$ *on* $\mathbf{E}$ *to* $\mathbf{E}_1$ *be of class* $C^{(n)}$ *uniformly on* $\mathbf{X}$ *is that the following conditions* (A) *and* (B) *be satisfied for* $k = n$.

(A) $F(x; \delta_1 x; \cdots; \delta_k x)$ *exists and is continuous in* x *uniformly* (*in the ordinary sense*) *with respect to its entire set of arguments for* $\| \delta_i x \| \leq 1$, $i = 1, \cdots, k$ *and* x *in* $\mathbf{X}$.

(B) *There exists a constant* M_k *with*

---

[19] We are indebted to the referee for helpful suggestions in this connection. Originally we employed the direct method, without using the results of Hildebrandt and Graves. For the general implicit function theorems see *Transactions of the American Mathematical Society*, vol. 29 (1927), pp. 127-153. From now on this paper will be referred to as HG.

[20] For the denition of functions of class $C^{(n)}$ see HG, p. 137 and 140.

$$\| F(x;\delta_1 x;\cdots;\delta_k x)\| \leq M_k \|\delta_1 x\| \cdots \|\delta_k x\|$$

for all x *in* $\mathbf{X}$.

Proof. The necessity of the conditions of the lemma is obvious from the definition of functions of class $C^{(n)}$ uniformly on $\mathbf{X}$ (uniformly $(\mathbf{X};1)$ in the notation of HG). To prove the sufficiency we note that condition (3), p. 137 of HG is automatically fulfilled since clearly

$$F(x) - F(x_0) = \int_0^1 F(x_0 + \sigma(x - x_0);\, x - x_0)\, d\sigma \tag{2.4}$$

for any two points x_0, x of $\mathbf{X}$. We now make use of the similar formula

$$\begin{aligned} F(x;\delta_1 x;\cdots;\delta_r x) &- F(x_0;\delta_1 x;\cdots;\delta_r x) \\ &= \int_0^1 F(x_0 + \sigma(x - x_0);\delta_1 x;\cdots;\delta_r x;\, x - x_0)\, d\sigma \\ &\qquad (x_0, x \text{ in } \mathbf{X}) \end{aligned} \tag{2.5}$$

to prove that conditions (A) and (B) are satisfied for $k = 1, 2, \cdots, n-1$, and therefore that the lemma holds. In fact (2.5) shows that (B) for $k = r + 1$ implies (A) for $k = r$. For each x_0 in $\mathbf{X}$, $F(x_0;\delta_1 x;\cdots;\delta_r x)$ is a multilinear function [21] in $\delta_1 x, \cdots, \delta_r x$ so that

$$\| F(x_0;\delta_1 x;\cdots;\delta_r x)\| \leq M_r(x_0) \|\delta_1 x\| \cdots \|\delta_r x\|.$$

Thus by (2.5) we see that (B) for $k = r + 1$ implies (B) for $k = r$. The lemma now follows by repeated applications of this argument.

The following theorem on existence and differentiability is independent of Theorems 1 and 2.

THEOREM 3. *Let* $\mathbf{X}$, $\mathbf{\Xi}$ *stand for the neighborhoods* $(\bar{x})_b$ *and* $(0)_\nu$ *of* $\mathbf{E}$ *and let the function* $H(x,\xi)$ *on* $\mathbf{XE}$ *to* $\mathbf{E}$ *have the following properties:*

(i) $H(x,\xi)$ *is of class* $C^{(n)}$ *in* (x,ξ) *uniformly on* $\mathbf{X\Xi}$ $(n \geq 1)$;

(ii) $H(x,\xi)$ *is homogeneous in* ξ *of degree* $r > 1$. *Then having chosen two real numbers* t_0, t_1, *there exist positive numbers* $\beta \leq b$ *and* $\lambda \leq \nu$ *such that*

(I) *for every* x_0, x_1 in $(\bar{x})_\beta$ *there is a solution* $x = \phi(t, x_0, x_1)$ *of the differential system*

$$d^2x/dt^2 = H(x, dx/dt), \qquad x(t_0) = x_0, \quad x(t_1) = x_1 \tag{2.6}$$

and this is the only solution with these end points satisfying

$$\| x(t) - \bar{x} \| < \lambda, \qquad \| dx/dt \| < \lambda$$

[21] A. D. Michal, *Annali di Matematica, loc. cit.*

(II) *the function $\phi(t, x_0, x_1)$ is of class $C^{(n)}$ in (x_0, x_1) uniformly on* $\boldsymbol{E}^2((\bar{x})_\beta)$.

Proof. As in the proof of Theorem 2 we deal with the following system of integral equations which is equivalent to the differential system (2. 6.):

$$(2.7) \qquad \begin{cases} x(t) = \int_0^c g(s,t)H(x(s),\xi(s))\,ds + x_0 + (t/c)(x_1 - x_0) \\ \xi(t) = \int_0^c g_t(s,t)H(x(s),\xi(s))\,ds + (1/c)(x_1 - x_0) \end{cases}$$

where we have chosen $t_0 = 0$, $t_1 = c$ for simplicity.

Let $\boldsymbol{E}_1$ be the space of pairs $y \equiv (x, \xi)$ where $x, y \in \boldsymbol{E}$, with the norm $\| y \| =$ greater of $\| x \|$, $\| \xi \|$, and let $\boldsymbol{E}_2$ be the space of continuous functions $Y \equiv y(t) \equiv (x(t), \xi(t))$ on $(0, c)$ to $\boldsymbol{E}_1$ with the norm $\| Y \| = \max_{0 \leq t \leq c} \| y(t) \|$. With the usual definitions of equality, addition and multiplication by reals, $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ are Banach spaces. Put $\omega \equiv (x_0, x_1)$. Next let $F_1(\omega, Y)$, $F_2(\omega, Y)$ respectively stand for the right members of the integral equations (2. 7) and write $F(\omega, Y) = (F_1(\omega, Y), F_2(\omega, Y))$, so that $F(\omega, Y)$ is a function with values in $\boldsymbol{E}_2$ for $\omega \in \boldsymbol{E}_1$, $Y \in \boldsymbol{E}_2$. Then the system (2. 7) may be replaced by the single equation

$$(2.8) \qquad G(\omega, Y) \equiv Y - F(\omega, Y) = 0.$$

We shall now show that hypotheses (H_1), (H_2), (H_3) of Theorem 4 of HG, p. 150 are satisfied by the functional equation (2. 8) with a proper choice of regions. In fact let $\boldsymbol{Y}_1$ be the region $\boldsymbol{X}\Xi$ of the composite space $\boldsymbol{E}_1$ and let $\boldsymbol{Y}_2$ be that subset of $\boldsymbol{E}_2$ consisting of all continuous functions $Y \equiv y(t)$ for which $y(t) \in \boldsymbol{Y}_1$, when $0 \leq t \leq c$. It is clear from its definition that $\boldsymbol{Y}_2$ is a bounded convex region. Take ω_0 as the point $x = \bar{x}$, $\xi = \bar{x}$ of $\boldsymbol{E}_1$ and Y_0 as the point $x(t) = \bar{x}$, $\xi(t) = 0$ of $\boldsymbol{E}_2$. If $\boldsymbol{\Omega}_1$ is any chosen bounded convex region of $\boldsymbol{E}_1$ containing ω_0, take the region $\boldsymbol{W}_0$ of the composite space $\boldsymbol{E}_1\boldsymbol{E}_2$ as $\boldsymbol{\Omega}_1\boldsymbol{Y}_2$, so that $\boldsymbol{W}_0$ is bounded, convex and contains the point (ω_0, Y_0).

(H_1) Since $H(x, \xi)$ is homogeneous in ξ it is clear that $F_1(\omega_0, Y_0) = \bar{x}$, $F_2(\omega_0, Y_0) = 0$, so that (ω_0, Y_0) is a solution of (2. 8).

(H_2) It is sufficient to show that $F_1(\omega, Y)$, $F_2(\omega, Y)$ are of class $C^{(n)}$ uniformly on $\boldsymbol{W}_0$. We give the details of proof for F_1 only, since those for F_2 are similar. Now $F_1(\omega, Y)$ can be written in the form

$$\Phi(Y) + \Lambda(\omega),$$

where Λ is a uniformly continuous linear function of ω on the bounded

region $\boldsymbol{\Omega}_1$, so that $F_1(\omega, Y)$ is of class $C^{(n)}$ uniformly on $\boldsymbol{W}_0$ if $\Phi(Y)$ has this property on $\boldsymbol{Y}_2$. Evidently

$$\Phi(Y) = \int_0^c g(s,t) H(x(s), \xi(s)) ds$$

is linear in H, while H may be considered as a function on $\boldsymbol{E}_2$ to $\boldsymbol{E}_2$ and written $H(Y)$. As a consequence of the uniformity conditions in (i) of our theorem it is easy to demonstrate that the differentials of $H(Y)$ up to and including the n-th exist for $Y \epsilon \boldsymbol{Y}_2$. An application of Lemma 3 then shows that $H(Y)$ is actually of class $C^{(n)}$ uniformly on $\boldsymbol{Y}_2$. The function Φ being linear in $H(Y)$, is also of class $C^{(n)}$ uniformly on $\boldsymbol{Y}_2$. Consequently $F_1(\omega, Y)$, $F_2(\omega, Y)$ and $F(\omega, Y)$ are of class $C^{(n)}$ uniformly on $\boldsymbol{W}_0 = \boldsymbol{\Omega}_1 \boldsymbol{Y}_2$.

(H_3) $$G(\omega_0, Y_0; \delta Y) = \delta Y - F(\omega_0, Y_0; \delta Y).$$

It is evident from the above discusssion that

$$F_1(\omega, Y; \delta Y) = \int_0^c g(s,t) H(x(s), \xi(s); \delta x(s), \delta\xi(s)) ds$$
$$F_2(\omega, Y; \delta Y) = \int_0^c g_t(s,t) H(x(s), \xi(s); \delta x(s), \delta\xi(s)) ds.$$

Since $H(x, \xi)$ is homogeneous of degree > 1 in ξ, the differential

$$H(x(s), \xi(s); \delta x(s), \delta\xi(s)) = H(x(s), \xi(s); \delta x(s)) + H(x(s), \xi(s); \delta\xi(s))$$

vanishes for $\xi(s) \equiv 0$ as is easy to see by calculating $H(x, \xi; \delta x)$ and $H(x, \xi; \delta\xi)$ by means of Gateaux's limit method. Hence for all δY, $G(\omega_0, Y_0; \delta Y) = \delta Y$ and (H_3) of Theorem 4 of HG is satisfied.

The hypotheses in Theorem 4 of HG hold. Hence by an application of this theorem the reader can readily complete the proof of our Theorem 3.

On differentiating the first equation of (2.7), we obtain with the aid of a theorem on total differentials proved by one of us [22] the following corollary.

Corollary 1. *The differentials* $\phi(t, w; \delta_1\omega; \cdots; \delta_k\omega)$ $(k = 1, 2, \cdots, n)$ *are uniformly modular* (*Cf. Lemma* 3, *property* (B)) *and continuous in* (t, ω) *uniformly with respect to their arguments for* $t \epsilon \boldsymbol{T}: t_1 < t < t_2$, $\omega \epsilon \boldsymbol{E}^2((\bar{x})_\beta)$ *and* $\| \delta_i\omega \| \leq 1$ $(i = 1, \cdots, k)$.

Corollary 2. *If in Theorem* 3, $n = 1$, *then* $\phi(t, \omega)$ *is of class* $C^{(1)}$ *in* (t, ω) *uniformly on* $\boldsymbol{TE}^2((\bar{x})_\beta)$.

[22] A. D. Michal, *Annali di Matematica, loc. cit.*

Proof. Since the pair of functions ϕ, $\partial\phi/\partial t$ satisfies (2.7) it can be shown by a simple calculation that $\partial\phi(t,\omega)/\partial t$ and $\phi(t,\omega;\delta\omega)$ are continuous in (t,ω) uniformly in $(t,\omega,\delta\omega)$ for $t\in\boldsymbol{T}$, $\omega\in\boldsymbol{E}^2((\bar{x})_\beta)$, $\|\delta\omega\|=1$, and have bounded norms for the same ranges of the variables. With the aid of Corollary 1 we obtain for a preassigned $\epsilon>0$ the inequality

$$\|\phi(t+\delta t,\omega+\delta\omega)-\phi(t,\omega)-\frac{\partial\phi(t,\omega)}{\partial t}\delta t-\phi(t,\omega;\delta\omega)\|$$
$$\leq|\delta t|\left\|\int_0^1\left\{\frac{\partial\phi(t+\sigma\delta t,\omega+\delta\omega)}{\partial t}-\frac{\partial\phi(t,\omega)}{\partial t}\right\}d\sigma\right\|+\epsilon\|\delta\omega\|$$

for t, $t+\delta t\in\boldsymbol{T}$, $\omega\in\boldsymbol{E}^2((\bar{x})_\beta)$ and $\|\delta\omega\|<\delta(\epsilon)$ ($\delta(\epsilon)$ independent of t and ω).

We return now to our study of the differential system (2.1).

THEOREM 4. *Suppose that besides satisfying hypotheses* (i), (ii) *of Lemma 2, the affine connection* $\Gamma(x,\xi,\eta)$ *has an n-th partial Fréchet differential* $\Gamma(x,\xi,\eta;\delta_1x;\cdots;\delta_nx)$ $n\geq 2$, *which is continuous in* x *for* $x\in(\bar{x})_{\bar{a}}$ *and for all* ξ,η. *Then there exist positive numbers* λ *and* d *such that the unique solution* $x(t,x_0,x_1)$ *of the system consisting of* (2.1) *and*

$$\|x(t,x_0,x_1)-\bar{x}\|\leq\bar{a},\quad\left\|\frac{\partial x(t,x_0,x_1)}{\partial t}\right\|\leq\lambda$$

has Fréchet differentials of orders $1,2,\cdots,n-1$ *in the set* (x_0,x_1) *which are uniformly continuous in* (t,x_0,x_1) *on* $\boldsymbol{TE}^2((\bar{x})_d)$.

Proof. As in Lemma 2 one can deduce that $\Gamma(x,\xi,\eta;\delta_1x;\cdots;\delta_kx)$ $(k=1,2,\cdots,n)$ is multilinear in $(\xi,\eta,\delta_1x,\cdots,\delta_kx)$, so that constants $b\leq\bar{a}$ and P_k exist such that

$$\|\Gamma(x,\xi,\eta;\delta_1x;\cdots;\delta_kx)\|\leq P_k\|\xi\|\,\|\eta\|\,\|\delta_1x\|\cdots\|\delta_kx\|$$

for $x\in(\bar{x})_b$ and hence, on using the integral formula for the difference repeatedly, that $\Gamma(x,\xi,\eta;\delta_1x;\cdots;\delta_lx)$ $(l=1,2,\cdots,n-1)$ is continuous in (x,ξ,η) uniformly with respect to its entire set of arguments for $x\in(\bar{x})_b$ and $\xi,\eta,\delta_1x,\cdots,\delta_lx$ each in norm ≤ 1. A formula proved by one of us [23] for the total differential of a bilinear function depending on a parameter gives

$$(2.9)\qquad \Gamma(x,\xi,\xi;\delta x,\delta\xi)=\Gamma(x,\xi,\xi;\delta x)+\Gamma(x,\xi,\delta\xi)+\Gamma(x,\delta\xi,\xi).$$

Repeated applications of this formula show that the $n-1$-st total differential of $\Gamma(x,\xi,\xi)$ exists. Each term of this total differential has properties of continuity and modularity similar to those of $\Gamma(x,\xi,\eta;\delta_1x;\cdots;\delta_lx)$, and there-

[23] A. D. Michal, *Annali di Matematica, loc. cit.*

fore by Lemma 3 $\Gamma(x, \xi, \xi)$ is of class $C^{(n-1)}$ uniformly on $\boldsymbol{E}((\bar{x})_b)\boldsymbol{E}^2((0)_\nu)$ where ν is any number ≤ 1. Hence by Theorem 3 there exist positive numbers $\beta(\nu)$ and $\lambda(\nu)$ such that $x(t, x_0, x_1)$ is of class $C^{(n-1)}$ in (x_0, x_1) uniformly on $\boldsymbol{E}^2((\bar{x})_{\beta(\nu)})$ and is unique among the class of functions for which

$$\| x(t, x_0, x_1) - \bar{x} \| < \lambda(\nu), \quad \left\| \frac{\partial x(t, x_0, x_1)}{\partial t} \right\| < \lambda(\nu).$$

Theorem 4 now follows by choosing ν small enough so that $\lambda(\nu)$ satisfies the hypotheses of Theorem 2, and by applying Corollary 1 of Theorem 3.

On making special use of Theorem 4 and Corollary 2 of Theorem 3 we obtain the following additional result on geodesics.

THEOREM 5. *Suppose that besides satisfying the hypotheses of Lemma 2, the affine connection $\Gamma(x, \xi, \eta)$ has a second partial Fréchet differential*[24] *$\Gamma(x, \xi, \eta; \delta_1 x; \delta_2 x)$ which is continuous in x for $x \in (\bar{x})_{\bar{a}}$ and for all ξ, η. Then there exist positive numbers λ and d such that the solution $x(t, x_0, x_1)$ of the system consisting of* (2.1) *and*

$$\| x(t, x_0, x_1) - \bar{x} \| \leq \bar{a}, \qquad \| dx/dt \| \leq \lambda$$

is of class $C^{(1)}$ uniformly on $\boldsymbol{TE}^2((\bar{x})_d)$.

If the real Banach space $\boldsymbol{E}$ is replaced by a complex Banach space,[25] i. e., a complete normed vector space closed under multiplication by complex numbers, Theorems 3, 4 and 5 are still valid. Thus, by a result on abstract analytic functions given independently by L. M. Graves and A. E. Taylor,[26] the solution $x(t, x_0, x_1)$ is analytic, and hence indefinitely Fréchet differentiable.

The attention of the reader is drawn to the fact that the validity of Theorems 2 and 5 does not depend on a particular parameterization of the geodesic $x(t, x_0, x_1)$. More precisely, the region $(\bar{x})_\delta$ is the same for all affine parameterizations.

---

[24] The conclusions of Theorem 5 will still be valid if we merely require that the first partial differential $\Gamma(x, \xi, \eta; \delta x)$ be continuous in (x, ξ, η) uniformly for $x \in (\bar{x})_b$, $\xi, \eta \in (0)_\nu$ and $\| \delta x \| \leq 1$. A similar remark of course holds for Theorem 4.

[25] By a linear function in a complex Banach space we shall mean a homogeneous function of degree one which is additive and continuous. It is not possible to prove homogeneity from additivity and continuity alone as is the case for real Banach spaces.

[26] A. E. Taylor, *Bulletin of the American Mathematical Society*, November (1935); L. M. Graves, *Bulletin of the American Mathematical Society*, October (1935). For the earlier work on abstract analytic function theory see R. S. Martin, *California Institute thesis* (1932); A. D. Michal and R. S. Martin, *Journal de Mathématiques Pures et Appliquées*, vol. 13 (1934), pp. 69-91; also N. Wiener, *Fundamenta Mathematicae* (1923), for abstractly valid analytic functions of a complex variable.

The uniqueness of the path passing through two points has been shown only for the class of paths such that $\| dx/dt \| < \lambda$. It is however possible to prove a modified uniqueness theorem where no restriction is placed on dx/dt.

THEOREM 6. *There exist two neighborhoods $(\bar{x})_\sigma$, $(\bar{x})_\tau$, $(\sigma < \tau)$ such that a unique path joins any two points of $(\bar{x})_\sigma$ and lies wholly within $(\bar{x})_\tau$.*

Proof. The proof depends on the following lemma, given by Whitehead [27] for the classical case.

LEMMA 4. *There is a neighborhood $(\bar{x})_\tau$ $(\tau < \delta)$ which is simple,* i. e., *not more than one path joins any two points of $(\bar{x})_\tau$ and lies wholly within the neighborhood.*

Let $\mu(t) = \lambda(t)(t - t_0)$ where

$$\lambda(t) = \max_{t_0 \leq s \leq t} \| dx/ds \|,$$

and $\mu_0 = \lambda(t_1 - t_0)$ where λ is defined in Theorem 2; let a μ-path be one for which $\mu(t) < \mu_0$, $t_0 \leq t \leq t_1$. Then Theorem 2 states that there is a unique μ_0-path joining any two points of $(\bar{x})_\delta$. Having shown that

$$\| x(s, x_0, x_1) - x_0 \| \geq \mu(s) - \tfrac{1}{2} M \mu^2(s)$$

one can proceed exactly as in the last part of Whitehead's proof. But on taking $t_0 = 0$ for simplicity and applying Lemma 1 with

$$\Phi(t) = \Gamma[x(t, x_0, x_1), dx/dt, dx/dt], \quad c = s, \quad m = M\lambda^2(s)$$

we find that

$$\left\| \frac{dx(t, x_0, x_1)}{dt} \right\| \leq \frac{sM\lambda^2(s)}{2} + \frac{\| x(s, x_0, x_1) - x_0 \|}{s} \quad \text{for} \quad 0 \leq t \leq s \leq t_1.$$

Now $\| dx/dt \|$ is a real continuous function of t on $0 \leq t \leq s$ and attains its maximum $\lambda(s)$ on this interval so that the required inequality and the lemma follow.

Since $\tau < \bar{a}$, Theorem 2 is valid with $\bar{a}$ replaced by τ, and the proof of Theorem 6 is complete.

3. *A differential equation in a normed ring.* Let $\boldsymbol{E}$ be a Banach space in which there exists on $\boldsymbol{E}^2$ to $\boldsymbol{E}$ an associative bilinear function denoted by xy.

[27] J. H. C. Whitehead, *Quarterly Journal of Mathematics*, vol. 3 (1932), pp. 38-39.

For simplicity, let the modulus of this bilinear function be unity, i.e., $\| xy \| \leq \| x \| \; \| y \|$. We also require that there exist a unit element I in $\boldsymbol{E}$. Clearly the functions $x + y$ and xy satisfy the postulates for an abstract ring.[28]

If the affine connection is taken to be

$$\Gamma(x, \xi, \eta) = \xi\eta,$$

the corresponding differential system for the paths, passing through two points is

$$(3.1) \qquad \begin{cases} d^2x/dt^2 = (dx/dt)^2 \\ x(0) = x_0, \qquad x(t_1) = x_1. \end{cases}$$

It is perhaps worthy of note that this extremely simple affinely connected space is not in general "flat," since the curvature tensor [29]

$$B(x, \xi, \delta_1 x, \delta_2 x) = \xi\delta_1 x\delta_2 x - \xi\delta_2 x\delta_1 x$$

vanishes identically if and only if the ring is commutative.

The system (3.1) can be integrated by means of elementary functions. In fact define $ln(I + x)$ and e^x as the respective power series

$$\sum_{n=1}^{\infty} ((-1)^{(n-1)}/n)x^n \quad \text{when} \quad \| x \| < 1$$

and

$$I + \sum_{n=1}^{\infty} (x^n/n!).$$

Then we can show that for sufficiently small values of $\| x_0 - x_1 \|$, the system (3.1) has the solution

$$(3.2) \qquad \phi(t, x_0, x_1) = x_0 - ln[I - (t/t_1)(I - e^{x_0 - x_1})]$$

By using a method similar to that of J. von Neumann [30] in the case of a finite matric ring, it may be shown that the relation $ln(e^x) = x$ holds in our normed ring providing that $\| I - e^x \| < 1$. Hence if $\| I - e^{x - x_0} \| < 1$, the boundary conditions are satisfied by the function $\phi(t, x_0, x_1)$. The ordinary rules for differentiation of a power series can be extended to the case where the coefficients are elements of our normed ring. Moreover Cauchy's rule for the multiplication of absolutely convergent series holds here. Hence, with the aid of Lemmas 1 and 2 and the uniqueness proof of Theorem 1, we have

---

[28] Van der Waerden, *Moderne Algebra*, vol. 1 (Berlin, 1930), pp. 36-37.

[29] A. D. Michal, *Annali di Matematica*, *loc. cit.*

[30] *Mathematische Zeitschrift*, vol. 30 (1929), p. 9.

THEOREM 7. *The function $\phi(t, x_0, x_1)$ defined in (3.2) satisfies the differential system (3.1) for $0 \leq t \leq t_1$ and for x_0, x_1 such that*

$$\| I - e^{x_1 - x_0} \| < 1.$$

If moreover

$$\| x_0 - x_1 \| < \lambda t_1/3 < 1/3,$$

then $\phi(t, x_0, x_1)$ is the unique solution among the class of functions for which

$$\| dx/dt \| < \lambda.$$

Given the power series

$$f(x) = \sum_n a_n x^n,$$

where a_n, x are elements of our ring, define

$$h_n(x_1, x_2, \cdots, x_n) = \sum (a_n/n!) x_{i_1} x_{i_2} \cdots x_{i_n},$$

the summation extending over all permutations of $x_1, x_2, \cdots, x_n$. It may be proved that the r-th differential of $f(x)$ exists for x in the region of convergence of $\sum a_n x^n$ and is given by

$$f(x; \delta_1 x; \delta_2 x; \cdots; \delta_r x) = \sum_{n=r}^{\infty} (n!/r!) h_n(x, x, \cdots, x, \delta_1 x, \delta_2 x, \cdots, \delta_r x).$$

Consequently the total differentials of all orders of $\phi(t, x_0, x_1)$ exist for $\| I - e^{x_0 - x_1} \| < 1$.

THEOREM 8. *The solution $\phi(t, x_0, x_1)$ of the differential system (3.1) has total Fréchet differentials of all orders for $\| I - e^{x_0 - x_1} \| < 1$.*

Besides the well known finite matric examples of a normed ring, we may take the class of ordered pairs of real functions $(x(r, s), x(u))$ with suitably defined operations,[31] where $x(r, s)$ and $x(r)$ are continuous for r, s in (0.1). Then the following integro-differential system is an instance of system (3.1).

$$\begin{cases} \dfrac{\partial^2 x(r, s, t)}{\partial t^2} = \displaystyle\int_0^1 \frac{\partial x(r, p, t)}{\partial t} \frac{\partial x(p, s, t)}{\partial t} dp \\ \qquad\qquad + \dfrac{\partial x(r, s, t)}{\partial t} \dfrac{\partial x(r, t)}{\partial t} + \dfrac{\partial x(r, s, t)}{\partial t} \dfrac{\partial x(s, t)}{\partial t} \\ \dfrac{\partial^2 x(r, t)}{\partial t^2} = \left(\dfrac{\partial x(r, t)}{\partial t} \right)^2 \\ x(r, s, 0) = x_0(r, s), \qquad x(r, 0) = x_0(r) \\ x(r, s, t_1) = x_1(r, s), \qquad x(r, t_1) = x_1(r). \end{cases}$$

[31] Compare with Evans' and Volterra's functional algebra. See Evans, *loc. cit.*, and Volterra, *loc. cit.* See also A. D. Michal and R. S. Martin, *loc. cit.*

4. *An integro-differential system in a functional geometry.* In certain functional differential geometries [32] the following integro-differential system is the analogue of the equations for the paths in a finite dimensional affine geometry.

$$(4.1)\quad \begin{cases} \dfrac{\partial^2 x^i(s)}{\partial s^2} + \Gamma_{\alpha\beta}{}^i \dfrac{\partial x^\alpha}{\partial s}\dfrac{\partial x^\beta}{\partial s} + M_\alpha{}^i \left(\dfrac{\partial x^\alpha}{\partial s}\right)^2 + 2N_\alpha{}^i \dfrac{\partial x^i}{\partial s}\dfrac{\partial x^\alpha}{\partial s} + P^i \left(\dfrac{\partial x^i}{\partial s}\right)^2 = 0, \\ \Gamma_{\alpha\beta}{}^i = \Gamma_{\beta\alpha}{}^i, \qquad x^i(s_0) = x_0{}^i, \quad x^i(s_1) = x_1{}^i. \end{cases}$$

Take the Banach space E of Theorem 2 to be the space of real functions x^i continuous over the interval (σ, τ) with $\| x^i \| = \max | x^i |$.

To satisfy the hypotheses (i) and (ii) of Lemma 2, we require that the functionals Γ, M, N, P be bounded and have continuous values when i, α, β are in (σ, τ) and x^i satisfies $\max\limits_i | x^i - \bar{x}^i | < a$. Let E_2 and E_3 be the spaces of continuous functions of two and three variables respectively with each variable ranging over (σ, τ) and the norm taken as the maximum of the absolute values. Then $\Gamma_{\alpha\beta}{}^i$, for example, is a function on E to E_3. For hypothesis (iii) to be satisfied it is sufficient that $\Gamma_{\alpha\beta}{}^i$, $M_\alpha{}^i$, $N_\alpha{}^i$, P^i have continuous first Fréchet differentials provided we take $\Gamma(x, \xi, \eta)$ to be

$$\Gamma_{\alpha\beta}{}^i \xi^\alpha \eta^\beta + M_\alpha{}^i \xi^\alpha \eta^\alpha + N_\alpha{}^i \xi^i \eta^\alpha + N_\alpha{}^i \xi^\alpha \eta^i + P^i \xi^i \eta^i.$$

Theorem 2 can now be applied immediately to the integro-differential system (4.1).

If in addition to the above hypotheses on the functionals Γ, M, N, P, we require that the second or higher order differentials of these functionals exist and are continuous the hypotheses of Theorems 4 and 5 are also satisfied.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

[32] A. D. Michal, *Proceedings of the National Academy of Sciences*, vol. 16 (January, 1930), pp. 88-94; *American Journal of Mathematics, loc. cit.*

In equation (4.1), the indices are continuous real variables on the interval (σ, τ) and a repetition of a subscript and superscript in a term denotes integration over (σ, τ). The values of the functionals $\Gamma_{\alpha\beta}{}^i[x^\lambda]$, $M_\alpha{}^i[x^\lambda]$, $N_\alpha{}^i[x^\lambda]$, $P^i[x^\lambda]$ are real continuous functions of three, two, two and one real variables respectively.

UNIVERSAL HOMOLOGY GROUPS.[1]

By NORMAN E. STEENROD.

Introduction. The development of algebraic topology of the past ten years has been marked by a constant effort to extend its principal theorems to ever more general spaces. Outstanding in this respect has been the endeavor to extend the Alexander duality theorem, the theory of manifolds, and to bring within the framework the theory of dimensionality. These investigations have compelled the introduction of homology groups based on chains and cycles with coefficients in a general topological abelian group.

There have been many reasons to suspect that one does not need to consider all possible coefficient groups to obtain a complete set of homology relations in a space. Thus, in the case of a complex, it has been proved that it is sufficient to know the homology groups based on integer coefficients. These groups and an abelian group $\mathfrak{G}$ determine completely the homology groups based on chains with coefficients in $\mathfrak{G}$. We say therefore that the integers form a universal coefficient group for the homology theory of a complex.

There are examples to show that this theorem does not hold if "complex" is replaced by "compact metric space." Alexandroff [5][2] has raised the problem of finding a universal coefficient group for a compact metric space. He surmised that the group $\mathfrak{X}$ of real numbers reduced modulo 1 is such a group. A proof of this is contained in the present paper. In fact, using Čech's definition [8] of the homology groups, we shall prove that $\mathfrak{X}$ is universal for the homology theory of a general topological space. To express the result differently: *the use of other coefficient groups cannot lead to any new topological invariants.* It will also be shown that $\mathfrak{X}$ is universal for the homology theory of the infinite cycles of an infinite complex. All of these results hold for relative homology groups in the sense of Lefschetz [15].

There is a dual result. Using the procedure of Čech, one may define, in the manner of Alexander [1], the dual homology groups of a topological space. Then the group of integers (the character group of $\mathfrak{X}$) is a universal coefficient group for this homology theory. This generalizes in a sense a theorem of Čech [9] which asserts that the integers are universal for the homology theory of the finite cycles of an infinite complex.

[1] The results of this paper were communicated in a note in the *Proceedings of the National Academy*, vol. 21 (1935), pp. 482-484.

[2] The numbers in brackets refer to the bibliography at the end of the paper.

These results have the implication that, in any investigation into the relations between a space and its homology groups, one obtains a theory with a maximum content by using either dual homology groups with integer coefficients or ordinary homology groups with coefficients in $\mathfrak{X}$. The two procedures are entirely equivalent, for the two types of groups of the same dimension are character groups of one another (Pontrjagin [17]). Since the dual homology groups with integer coefficients are somewhat easier to handle, this result constitutes a strong argument for the future exclusive use of these groups.

In part I we formalize a well-known tool of the topologist: mapping systems and homomorphism systems. These notions are found first in the projection spectrum of Alexandroff [4], and they have since been used by Pontrjagin [16], Lefschetz [15], and others. In keeping with the ideas of Čech [8], we do not restrict ourselves to sequences but admit partially ordered systems. This requires a complete treatment of the subject. By using the van Kampen [13] extension of the character group theory of Pontrjagin [17], we establish two theorems on the representations of bicompact abelian groups as limit groups of homomorphism systems of finite dimensional, locally-connected groups.

In part II we consider first the homology groups of a finite complex. We give the structure of these groups in terms of the coefficient group and known invariants of the complex. A proof of this decomposition, omitting all continuity considerations, can be found in the book of Alexandroff and Hopf [7; pp. 228-240]. Then following Čech [8], we define the homology groups of a general topological space. We establish the decomposition of these groups into direct sums of "torsion groups" and "reduced homology groups." In this is contained a corresponding direct sum theorem for the dual homology groups.

Part III contains the proof proper that $\mathfrak{X}$ is a universal coefficient group for a topological space. In part IV we indicate the modifications necessary to obtain the same theorem for an infinite complex.

The first appendix contains a proof that a connected, bicompact, abelian group is continuously isomorphic with its 1-dimensional homology group over $\mathfrak{X}$. This indicates a probability that almost any bicompact abelian group can appear as a modulo 1 homology group of some space. In the second appendix we give a number of examples to support earlier statements of the paper.

This investigation was made under the guidance of Professor S. Lefschetz. I am indebted to him and to Professor A. W. Tucker for many helpful suggestions. I am likewise indebted to Dr. Leo Zippin for general criticisms

and for constant assistance with the group theoretic considerations which, it is clear, occur throughout.

Certain results of this paper were obtained independently by Čech [10]. During the past year the author has had the benefit of many conversations with Professor Čech. These have contributed materially to the generality of our results.

I. Homomorphism Systems.

1. Definitions and notation. We shall have occasion to use only commutative groups; we may therefore, without confusion, omit the word commutative. All groups are written additively. Groups are denoted by capital German letters and their elements by small German letters.

By *topological space* we mean a set in which certain subsets are called open so that the axioms 1 to 4 of Hausdorff [12, p. 228] are satisfied. In particular we use only the separation axiom which asserts that a point is closed. Spaces and their elements are denoted by Latin letters.

A *bicompact space* is a topological space with the property that every covering by open sets contains a finite covering. A closed subset of a bicompact space is bicompact;[3] and a continuous image of a bicompact space is bicompact. By interpreting the definition of bicompactness in terms of the complementary closed sets, one obtains the following characterization:

LEMMA 1.1. *For a space to be bicompact it is necessary and sufficient that, for any collection $\{F^a\}$ of closed subsets, $\prod F^a \neq 0$ if this relation holds for each finite subcollection of $\{F^a\}$.*

A *topological group* is a group whose elements are the points of a topological space whose neighborhoods satisfy the additional axiom:

α) if U is a neighborhood of $\mathfrak{x} - \mathfrak{y}$ there are neighborhoods V and W of $\mathfrak{x}$ and $\mathfrak{y}$ respecitvely such that $\mathfrak{x}' - \mathfrak{y}' \in U$ whenever $\mathfrak{x}' \in V$ and $\mathfrak{y}' \in W$.

Since the zero point is closed, it follows from α) that a topological group satisfies the separation axiom 5 of Hausdorff.

By setting $\mathfrak{x} = 0$ in α), we find that $f(\mathfrak{y}) = -\mathfrak{y}$ is a continuous function of $\mathfrak{y}$. Combining this with α) we obtain that $f(\mathfrak{x}, \mathfrak{y}) = \mathfrak{x} + \mathfrak{y}$ is continuous simultaneously in $\mathfrak{x}$ and $\mathfrak{y}$. In general $\sum_{i=1}^{n} a_i \mathfrak{x}_i$ ($a_i =$ integer) is continuous simultaneously in all its variables.

If $\mathfrak{A}$ and $\mathfrak{B}$ are topological groups and H is a homomorphic mapping of

[3] For proofs of these statements and the following lemma see Alexandroff and Urysohn [6]. They assume that a bicompact space has the separation axiom 5 of Hausdorff, but this is not essential.

$\mathfrak{A}$ into $\mathfrak{B}$ which is continuous at the zero of $\mathfrak{A}$, then, by a simple application of α), H is continuous at each point of $\mathfrak{A}$. Thus to prove the continuity of a homomorphism it suffices to prove its continuity at the zero.

If $\mathfrak{B}$ is a subgroup of $\mathfrak{A}$, the residue (factor) group $\mathfrak{A}$ mod $\mathfrak{B}$ is denoted $\mathfrak{A} - \mathfrak{B}$. If $\mathfrak{A}$ is a topological group and $\mathfrak{B}$ is closed in $\mathfrak{A}$, we may designate as neighborhoods in $\mathfrak{A} - \mathfrak{B}$ the images of neighborhoods of $\mathfrak{A}$. Then, since $\mathfrak{B}$ is closed, it can be proved that $\mathfrak{A} - \mathfrak{B}$ is a topological group.

A topological group in which each point is a neighborhood of itself is called a *discrete group.*

Let $\{\mathfrak{A}^\alpha\}$ be a well-ordered collection of topological groups. By the *direct sum* $\sum \mathfrak{A}^\alpha$ we mean the topological group whose elements are collections $\{\mathfrak{a}^\alpha\}$, $\mathfrak{a}^\alpha \in \mathfrak{A}^\alpha$, in which difference is defined by $\{\mathfrak{a}^\alpha\} - \{\mathfrak{a}'^\alpha\} = \{\mathfrak{a}^\alpha - \mathfrak{a}'^\alpha\}$, and a neighborhood of $\{\mathfrak{a}^\alpha\}$ is defined by a finite number of indices α_i $(i = 1, \cdots, k)$ and neighborhoods V^{α_i} of the coördinates $\mathfrak{a}^{\alpha_i}$ and consists of all points $\{\mathfrak{a}'^\alpha\}$ such that $\mathfrak{a}'^{\alpha_i} \in V^{\alpha_i}$ $(i = 1, \cdots, k)$. It is not difficult to prove

LEMMA 1.2. *If $\mathfrak{A} = \sum \mathfrak{A}^\alpha$ and the subgroup $\mathfrak{B}$ of $\mathfrak{A}$ is the direct sum $\sum \mathfrak{B}^\alpha$ where $\mathfrak{B}^\alpha$ is a closed subgroup of $\mathfrak{A}^\alpha$, then $\mathfrak{A} - \mathfrak{B}$ and $\sum (\mathfrak{A}^\alpha - \mathfrak{B}^\alpha)$ are bicontinuously isomorphic.*

2. Inverse mapping systems. Let $\{A^\alpha\}$ be a collection of topological spaces (the index α runs over the set of ordinals less than a fixed but arbitrary ordinal κ). Suppose that to some pairs (α, β) of ordinals $< \kappa$ there corresponds a continuous mapping $M_\beta{}^\alpha$ of A^β into A^α, so that the following conditions are satisfied:

a) $M_\alpha{}^\alpha$, for each α, is defined and is the identity,

b) if $M_\gamma{}^\beta$ and $M_\beta{}^\alpha$ are defined, then $M_\gamma{}^\alpha$ is defined and is the product of the two transformations: $M_\gamma{}^\alpha = M_\beta{}^\alpha M_\gamma{}^\beta$,

c) to each pair (α, β) of ordinals $< \kappa$ corresponds a γ such that $M_\gamma{}^\alpha$ and $M_\gamma{}^\beta$ are defined.

Such a collection of spaces and continuous mappings we shall call an *inverse system.*

The existence of $M_\beta{}^\alpha$ we shall indicate by $A^\beta \to A^\alpha$; $M_\beta{}^\alpha$ is called a *projection* of the system. The space A^γ postulated in c) is called a *common refinement* of A^α and A^β. Iteration of c) shows that any finite number of spaces of the system have a common refinement.

We define the limit space A of an inverse system S. The points of A shall be all collections $\{a^\alpha\}$ consisting of one point a^α from each A^α such that, if $A^\beta \to A^\alpha$, then $M_\beta{}^\alpha(a^\beta) = a^\alpha$. The point a^α is said to be the coördinate in

A^α of the point $a = \{a^\alpha\}$ of A. If V^β is a neighborhood of the fixed coördinate a^β of a, the set V in A of all points $\tilde{a} = \{\tilde{a}^\alpha\}$ satisfying $\tilde{a}^\beta \epsilon V^\beta$ is said to be a neighborhood of a in A. Allowing V^β to range over a complete set of neighborhoods of a^β and β to range over all ordinals $< \kappa$, we define in this way a complete set of neighborhood of a in A. We must verify that these neighborhoods make of A a topological space.

It is trivial that $a \epsilon V$. Suppose $\tilde{a} \epsilon V$; then let $\tilde{V}^\beta$ be a neighborhood of $\tilde{a}^\beta$ lying in V^β. It follows that $\tilde{V}^\beta$ determines a neighborhood $\tilde{V}$ of $\tilde{a}$ contained in V. Suppose V and V' are two neighborhoods of a determined by V^α and V^β respectively. Let A^γ be a common refinement of A^α and A^β. As $M_\gamma{}^\alpha$ and $M_\gamma{}^\beta$ are continuous, there is a neighborhood V^γ of a^γ such that $M_\gamma{}^\alpha(V^\gamma) \subset V^\alpha$ and $M_\gamma{}^\beta(V^\gamma) \subset V^\beta$. Then V^γ determines a neighborhood V'' of a contained in V and V'. Suppose a and b are distinct points of A. Then, for some β, their coördinates a^β and b^β are distinct. Then a neighbor-V^β of a^β not containing b^β determines a neighborhood V of a not containing b. This proves that A is a topological space. We remark that, if in each A^α, two points can be separated by neighborhoods, the same holds true in A.

The mapping $M^\alpha(a) = a^\alpha$ of A into A^α is continuous and is called the projection of A into A^α. If A and the projections M^α are added to S, we obtain a new inverse system whose limit space is A.

If each A^α is regular, A is regular. If the system S is countable, then the first (second) countability axiom in each A^α implies the first (second) countability axiom in A. If, in each A^α, we substitute an equivalent system of neighborhoods, we obtain an equivalent system of neighborhoods in A.

As we shall be frequently concerned with inverse systems of bicompact spaces, we prove the following important lemma for such systems.

LEMMA 2.1. *If S is an inverse system of bicompact spaces, and, for a fixed δ, a^δ is a point of A^δ such that $A^\beta \rightarrow A^\delta$ implies $M_\beta{}^\delta(A^\beta) \supset a^\delta$, there is a point a of the limit space A such that $M^\delta(a) = a^\delta$.*

We suppose the spaces of S are well-ordered so that A^δ occurs as the first element (i. e. we suppose $\delta = 1$). The coördinates of the point a are chosen inductively. The first coördinate is a^1. Suppose we have chosen a^α for each $\alpha < \beta$ so that, if A^γ is a common refinement of $A^{\alpha_1}, \cdots, A^{\alpha_k}$ $(\alpha_i < \beta)$, the sets $F^{\alpha_1}, \cdots, F^{\alpha_k}$ of inverse images in A^γ of $a^{\alpha_1}, \cdots, a^{\alpha_k}$, respectively, have a non-vacuous intersection. Clearly the choice of a^1 satisfies this condition. Suppose $\alpha < \beta$ and $A^{\gamma(\alpha)}$ is a refinement of A^α and A^β. Denote by $F^{\gamma(\alpha)}$ the image in A^β of the inverse image in $A^{\gamma(\alpha)}$ of a^α. We assert that

$\prod_{\substack{\gamma(a)\\ a<\beta}} F^{\gamma(a)}$ is non-vacuous. As $F^{\gamma(a)}$ is closed, by Lemma 1.1, it is sufficient to prove that $\prod_{i=1}^{k} F^{\gamma(a_i)}$ is non-vacuous for any finite set $\gamma(\alpha_i)$ $(i=1,\cdots,k)$. Let A^γ be a refinement of $A^{\gamma(a_i)}$ $(i=1,\cdots,k)$. By the hypothesis of the induction, the inverse images in A^γ of $a^{a_1},\cdots,a^{a_k}$ have a non-vacuous intersection F^γ whose image in A^β lies in each $F^{\gamma(a_i)}$. Thus $\prod F^{\gamma(a_i)}$ is non-vacuous, and we may select the coördinate a^β of a in this set. The hypothesis of the induction is still satisfied. For, if A^γ is a refinement of $A^\beta, A^{a_1},\cdots, A^{a_n}$, we may find a common refinement A^ϵ of A^γ and $A^{\gamma(a_i)}$ $(i=1,\cdots,k)$. Since A^ϵ is a refinement of $A^{\gamma(a_i)}$, the projection on A^β of the inverse image of a^{a_i} in A^ϵ equals $F^{\gamma(a_i)}$. As $a^\beta \epsilon F^{\gamma(a_i)}$, the inverse images of a^β and a^{a_i} $(i=1,\cdots,k)$ in A^ϵ have a non-vacuous intersection. The projection on A^γ of this intersection is common to the inverse images in A^γ of a^β and a^{a_i} $(i=1,\cdots,k)$.

The procedure outlined determines a collection $\{a^a\}$ of one point from each space. If $A^\beta \rightarrow A^a$, since A^β is a common refinement of A^a and A^β, the inverse images in A^β of a^a and a^β have a non-vacuous intersection. Thus a^β lies in the inverse image of a^a. This means $M_\beta{}^a(a^\beta)=a^a$; and $\{a^a\}$ is the desired point of A.

THEOREM 2.1. *The limit space of an inverse system of bicompact spaces is non-vacuous and bicompact.*

If $A^\beta \rightarrow A^1$, let $A_1{}^\beta$ be the image of A^β in A^1. Now each $A_1{}^\beta$ is closed and any finite number of sets $A_1{}^{\beta_i}$ $(i=1,\cdots,k)$ have a non-vacuous intersection; for, if A^γ is a refinement of A^{β_i} $(i=1,\cdots,k)$, $A_1{}^\gamma$ lies in this intersection. Thus, by Lemma 1.1, $\prod A_1{}^\beta$ is non-vacuous; and there is a point a^1 of A^1 contained in the image of each A^β such that $A^\beta \rightarrow A^1$. By Lemma 2.1, there is a point a of A such that $M^1(a)=a^1$. Thus A is non-vacuous.

Let F be a closed set in A. We shall prove that $F^a = M^a(F)$ is closed in A^a. If $A^\beta \rightarrow A^a$, the statement $M_\beta{}^a(F^\beta)=F^a$ implies $M_\beta{}^a(\bar{F}^\beta)=\bar{F}^a$ (the bar denoting the closure). Suppose $\bar{F}^a - F^a$ contains a point a^a. By Lemma 2.1, the limit space of the inverse system $\{\bar{F}^\beta\}$ of bicompact spaces contains a point a whose projection in $\bar{F}^a$ is a^a. Clearly $a \epsilon A$. In each neighborhood of this point are points of F; hence $a \epsilon \bar{F} = F$. This implies $a^a \epsilon F^a$ which is impossible. Therefore $F^a = \bar{F}^a$.

To prove that A is bicompact we shall use the sufficiency of the condition of Lemma 1.1. Let $\{F_\lambda\}$ be any collection of closed sets in A such that any finite subcollection has a non-vacuous intersection. Let $F_\lambda{}^a$ be the projection of F_λ in A^a. As proved above, $F_\lambda{}^a$ is closed. Any finite number of the sets

$F_\lambda{}^\alpha$ (α fixed) have a non-vacuous intersection; for this intersection contains the projection in A^α of the intersection of the corresponding sets in A. Therefore $\prod_\lambda F_\lambda{}^\alpha = F^\alpha$ is non-vacuous. If $A^\beta \to A^\alpha$, it is obvious that $M_\beta{}^\alpha(F^\beta) \subset F^\alpha$. By the first part of the theorem, the inverse system consisting of the projections $\{M_\beta{}^\alpha\}$ and the bicompact spaces $\{F^\alpha\}$ has a non-vacuous limit set $F \subset A$. In each neighborhood of a point of F there are points of F_λ; hence $F \subset \bar{F}_\lambda = F_\lambda$. As λ is arbitrary, $F \subset \prod F_\lambda$; and the theorem is proved.

We shall need later the following lemma whose proof properly belongs here.

LEMMA 2.2. *If $S = \{A^\alpha\}$ is an inverse system of bicompact spaces and U is an open set of A^α containing the image $M^\alpha(A)$ of the limit space A of S, then there is a refinement A^β of A^α such that $M_\beta{}^\alpha(A^\beta) \subset U$.*

If, to the contrary, the image in A^α of each refinement A^β of A^α meets $A^\alpha - U$, the collection of closed sets consisting of $A^\alpha - U$ and these images will have a non-vacuous intersection (use Lemma 1.1). By Lemma 2.1, this intersection which is a subset of $A^\alpha - U$ contains the coördinate of a point of A. The contradiction proves the lemma.

3. Equivalence of inverse systems. If S_1 and S_2 are two inverse systems, and the spaces and mappings of S_2 are included among those of S_1, then S_2 is called a *subsystem* of S_1. A subsystem S_2 of S_1 is said to be *complete* if each space of S_1 has a refinement in S_2.

LEMMA 3.1. *If S_2 is a complete subsystem of S_1, the limit spaces of S_1 and S_2 are homeomorphic.*

The coördinates of a point of the limit space A_1 of S_1 in the spaces of S_2 are the coördinates of a point of the limit space A_2 of S_2. The mapping of A_1 into A_2 so defined is 1 : 1 and bicontinuous. First, the image of A_1 in A_2 covers A_2; for, any point a_2 of A_2 has a projection in each space of S_2 and therefore, since S_2 is complete in S_1, a unique projection in each space of S_1. The set of these projections are the coördinates of a point a_1 of A_1 whose image in A_2 is a_2. Next, if a_1 and a'_1 are distinct in A_1, their images in A_2 are distinct. The distinctness of a_1 and a'_1 implies that of their projections $a_1{}^\alpha$ and $a'_1{}^\alpha$ for some α. Let A^β be a refinement of A^α in S_2. Then $a_1{}^\beta$ and $a'_1{}^\beta$ are distinct; but this implies that a_2 and a'_2 are distinct. Finally, the correspondence is bicontinuous. For, if V_2 is a neighborhood of a_2 determined by V^β in A^β (A^β in S_2), the neighborhood V_1 of a_1 determined by V^β images into V_2. If V_1 is a neighborhood of a_1 determined by V^α in A^α, let V^β be a neighborhood of $a_1{}^\beta$ in the refinement A^β of A^α (A^β in S_2) whose image in A^α

lies in V^α. Then V^β determines a neighborhood V_2 of a_2 covered by the image of V_1.

Two inverse systems are said to be *immediately equivalent* if both are complete subsystems of a third. Now it is not yet clear that this relation is a transitive one; so we define the following clearly transitive relation. Two inverse systems S and S' are said to be *equivalent* if there exists a finite ordered set of inverse systems $S, S_1, \cdots, S_k, S'$ such that neighboring pairs are immediately equivalent. As an immediate consequence of this definition and Lemma 3.1, we have

THEOREM 3.1. *Equivalent inverse systems have homeomorphic limit spaces.*

A countable inverse system is called an *inverse sequence* if, for any pair of its spaces, at least one is a refinement of the other. A simple induction proves that any countable inverse system contains a complete subsequence.

4. Inverse homomorphism systems. An inverse system is said to be an *inverse homomorphism system* if each space of the system is a topological group and each mapping of the system is a continuous homomorphism.[4] The homomorphisms of the system are denoted by $H_\beta{}^\alpha$. In the expression "inverse homomorphism system" we shall omit the word "homomorphism" when it is clear that we are dealing with groups.

The limit space $\mathfrak{A}$ of the inverse system $S = \{\mathfrak{A}^\alpha\}$ is converted into a topological group by the following definition of addition. If $\mathfrak{a}_1 = \{\mathfrak{a}_1{}^\alpha\}$, $\mathfrak{a}_2 = \{\mathfrak{a}_2{}^\alpha\}$ are two points of $\mathfrak{A}$, then $\mathfrak{a}_1 + \mathfrak{a}_2$ is $\{\mathfrak{a}_1{}^\alpha + \mathfrak{a}_2{}^\alpha\}$ which is clearly an element of $\mathfrak{A}$. The zero of $\mathfrak{A}$ is $\{0\}$, and $-\{\mathfrak{a}^\alpha\} = \{-\mathfrak{a}^\alpha\}$. It is a simple matter to prove that $\mathfrak{A}$ is a group. We consider now the axiom α) of No. 1. Let U be a neighborhood of $\mathfrak{a}_1 - \mathfrak{a}_2$ determined by U^α in $\mathfrak{A}^\alpha$. As $\mathfrak{A}^\alpha$ is a topological group, there are neighborhoods V^α and W^α of $\mathfrak{a}_1{}^\alpha$ and $\mathfrak{a}_2{}^\alpha$, respectively, as postulated in α). Then V^α and W^α determine neighborhoods V and W in $\mathfrak{A}$ satisfying α). Thus $\mathfrak{A}$ is a topological group. Let us remark that $\mathfrak{a} \to \mathfrak{a}^\alpha$ is a homomorphism of $\mathfrak{A}$ into $\mathfrak{A}^\alpha$.

As in No. 3, we define for inverse homomorphism systems the notions of subsystem, complete subsystems, and equivalence. Noting that, for an inverse homomorphism system, the correspondence set up in Lemma 3.1 is an isomorphism, we have

[4] It is clear that a similar definition can be made for non-commutative groups, and the limit group described below exists.

THEOREM 4.1. *Equivalent inverse homomorphism systems have bicontinuously isomorphic limit groups.*

5. Direct homomorphism systems. Let $\{\mathfrak{B}^\alpha\}$ be a well-ordered collection of groups and, for some pairs $\mathfrak{B}^\alpha$, $\mathfrak{B}^\beta$ let there be defined a homomorphism $H^*{}_\beta{}^\alpha$ of $\mathfrak{B}^\beta$ into $\mathfrak{B}^\alpha$ so that axioms a) and b) of No. 2 are satisfied and, in place of c), the axiom

c') to each pair (α, β) corresponds a γ such that $H^*{}_\alpha{}^\gamma$ and $H^*{}_\beta{}^\gamma$ are defined.

Such a system S^* of groups and homomorphisms we shall call a *direct homomorphism system.*

We define the limit group $\mathfrak{B}$ of the direct system S^*. An element of $\mathfrak{B}$ will be a collection $\{\mathfrak{b}^\alpha\}$ of one element from each of a subcollection of the groups of S^* satisfying the condition that, if $\mathfrak{B}^\alpha \rightarrow \mathfrak{B}^\beta$ and $\mathfrak{B}^\alpha$ is in the subcollection, so also is $\mathfrak{B}^\beta$ and $H^*{}_\alpha{}^\beta(\mathfrak{b}^\alpha) = \mathfrak{b}^\beta$. Two elements of $\mathfrak{B}$ are equal if in some group of S^* their coördinates are defined and are equal. Let $\mathfrak{b}_1$ and $\mathfrak{b}_2$ be two elements of $\mathfrak{B}$. Consider the collection of groups of S^* in each of which the coördinates of both $\mathfrak{b}_1$ and $\mathfrak{b}_2$ are defined. In each group of this collection let us add the coördinates of $\mathfrak{b}_1$ and $\mathfrak{b}_2$. The sums so obtained determine an element of $\mathfrak{B}$ called the sum $\mathfrak{b}_1 + \mathfrak{b}_2$ of $\mathfrak{b}_1$ and $\mathfrak{b}_2$. It is verified without difficulty that $\mathfrak{B}$ constitutes a group with addition so defined.

If $\mathfrak{b}^\alpha$ is an element of $\mathfrak{B}^\alpha$ the collection of its images in each $\mathfrak{B}^\beta$ such that $\mathfrak{B}^\alpha \rightarrow \mathfrak{B}^\beta$ determines an element $\mathfrak{b}$ of $\mathfrak{B}$. The correspondence $H^*{}_\alpha(\mathfrak{b}^\alpha) = \mathfrak{b}$ so defined is a homomorphism.

We assign to $\mathfrak{B}$ the discrete topology. The justification will be found in the theorem of the next section. We might remark that, in case each $\mathfrak{B}^\alpha$ is a topological group and each $H^*{}_\beta{}^\alpha$ is continuous, the method of No. 2 defines neighborhoods in $\mathfrak{B}$ which in general fail to satisfy the axiom that two neighborhoods of a point contain a third in their common part. Other obvious methods fail to make of $\mathfrak{B}$ a topological group.

6. Dual systems. Let $\mathfrak{X}$ be the group of real numbers reduced modulo 1. Let S be an inverse system of bicompact groups $\{\mathfrak{A}^\alpha\}$. Let $\mathfrak{B}^\alpha$ be the discrete group of continuous characters [5] of $\mathfrak{A}^\alpha$. The element of $\mathfrak{X}$ determined by $\mathfrak{b}^\alpha \in \mathfrak{B}^\alpha$ and $\mathfrak{a}^\alpha \in \mathfrak{A}^\alpha$ we denote by $(\mathfrak{a}^\alpha, \mathfrak{b}^\alpha)$. Suppose $\mathfrak{A}^\beta \rightarrow \mathfrak{A}^\alpha$. The result of the homomorphisms $\mathfrak{A}^\beta \rightarrow \mathfrak{A}^\alpha$ then $\mathfrak{A}^\alpha \rightarrow \mathfrak{X}$ as determined by the element $\mathfrak{b}^\alpha \in \mathfrak{B}^\alpha$ is a continuous character $\mathfrak{b}^\beta$ of $\mathfrak{A}^\beta$. The correspondence $H^*{}_\alpha{}^\beta(\mathfrak{b}^\alpha) = \mathfrak{b}^\beta$

[5] A continuous character of $\mathfrak{A}$ is a continuous homomorphic mapping of $\mathfrak{A}$ into the group $\mathfrak{X}$ (see [17; p. 362]).

so obtained is a homomorphism of $\mathfrak{B}^\alpha$ into a subgroup of $\mathfrak{B}^\beta$. This homomorphism is defined to satisfy

$$(6.1) \qquad (\mathfrak{a}^\beta, H^*{}_\alpha{}^\beta(\mathfrak{b}^\alpha)) = (H_\beta{}^\alpha(\mathfrak{a}^\beta), \mathfrak{b}^\alpha)$$

for all $\mathfrak{a}^\beta \epsilon \mathfrak{A}^\beta$ and $\mathfrak{b}^\alpha \epsilon \mathfrak{B}^\alpha$. The collection of groups $\{\mathfrak{B}^\alpha\}$ and homomorphisms $\{H^*{}_\beta{}^\alpha\}$ form a direct system S^* called the system dual to S.

Conversely, given a direct system S^* of discrete groups, by passing to character groups we can define an inverse system S of bicompact groups having S^* as its dual system.

THEOREM 6.1. *If S is an inverse system of bicompact groups and S^* is its dual direct system, then the limit group of S^* is the group of all continuous characters of the limit group of S.*

Let $\mathfrak{b}$ be an arbitrary element of the limit group $\mathfrak{B}$ of S^*. The function

$$(6.2) \qquad \mathfrak{b}(\mathfrak{a}) = (H^\alpha(\mathfrak{a}), \mathfrak{b}^\alpha)$$

where $\mathfrak{b}^\alpha$ is some fixed coördinate of $\mathfrak{b}$ is a continuous character of the limit group $\mathfrak{A}$ of S. By (6.1), if $\mathfrak{B}^\alpha \rightarrow \mathfrak{B}^\beta$,

$$(H^\alpha(\mathfrak{a}), \mathfrak{b}^\alpha) = (H_\beta{}^\alpha H^\beta(\mathfrak{a}), \mathfrak{b}^\alpha) = (H^\beta(\mathfrak{a}), H^*{}_\alpha{}^\beta(\mathfrak{b}^\alpha)) = (H^\beta(\mathfrak{a}), \mathfrak{b}^\beta).$$

Therefore the character defined by (6.2) is independent of the particular coördinate $\mathfrak{b}^\alpha$ chosen; so the character $\mathfrak{b}(\mathfrak{a})$ may be uniquely associated with the element $\mathfrak{b}$ of $\mathfrak{B}$. In this way $\mathfrak{B}$ is a group of continuous characters of $\mathfrak{A}$.

Let $\mathfrak{a}$ be a non-zero element of $\mathfrak{A}$. Then, for some α, the coördinate $\mathfrak{a}^\alpha$ of $\mathfrak{a}$ is not equal to zero. Let $\mathfrak{b}^\alpha$ be a character of $\mathfrak{A}^\alpha$ such that $(\mathfrak{a}^\alpha, \mathfrak{b}^\alpha) \neq 0$. The image $\mathfrak{b}$ of $\mathfrak{b}^\alpha$ in $\mathfrak{B}$ is then a character of $\mathfrak{A}$ such that $\mathfrak{b}(\mathfrak{a}) \neq 0$. Hence the annihilator [6] of $\mathfrak{B}$ in $\mathfrak{A}$ is the zero.

Let $\mathfrak{b}$ be a non-zero element of $\mathfrak{B}$. Then each coördinate $\mathfrak{b}^\alpha$ of $\mathfrak{b}$ is not equal to zero. Let $\tilde{\mathfrak{A}}^\alpha$ be the annihilator of $\mathfrak{b}^\alpha$ in $\mathfrak{A}^\alpha$. Let U be a neighborhood of zero in $\mathfrak{X}$ containing no subgroup of $\mathfrak{X}$ other than the zero.[7] As $\mathfrak{b}^\alpha$ maps $\tilde{\mathfrak{A}}^\alpha$ into zero, there is a neighborhood V of $\tilde{\mathfrak{A}}^\alpha$ which is mapped by $\mathfrak{b}^\alpha$ into U. Suppose, as is impossible, that $H^\alpha(\mathfrak{A}) \subset \tilde{\mathfrak{A}}^\alpha$ for a fixed $\mathfrak{a}$. By Lemma 2.2, there is a refinement $\mathfrak{A}^\beta$ of $\mathfrak{A}^\alpha$ such that $H_\beta{}^\alpha(\mathfrak{A}^\beta) \subset V$. Then the coördinate $\mathfrak{b}^\beta$ of $\mathfrak{b}$ is defined, and, as $\mathfrak{b}^\beta \neq 0$, there is an element $\mathfrak{a}^\beta$ of $\mathfrak{A}^\beta$ such that $(\mathfrak{a}^\beta, \mathfrak{b}^\beta) \neq 0$. However

$$(\mathfrak{a}^\beta, \mathfrak{b}^\beta) = (\mathfrak{a}^\beta, H^*{}_\alpha{}^\beta(b^\alpha)) = (H_\beta{}^\alpha(\mathfrak{a}^\beta), \mathfrak{b}^\alpha) \neq 0;$$

[6] That is, the set of elements of $\mathfrak{A}$ mapping each element of $\mathfrak{B}$ into zero.

[7] The complement of the closed interval [1/3, 2/3] is such a neighborhood.

so there is an element $\mathfrak{a}^\alpha$ of $H_\beta{}^\alpha(\mathfrak{A}^\beta)$ such that $(\mathfrak{a}^\alpha, \mathfrak{b}^\alpha) \neq 0$. But $H_\beta{}^\alpha(\mathfrak{A}^\beta)$ is a closed subgroup of V; and, as $\mathfrak{b}^\alpha$ maps V into U, $\mathfrak{b}^\alpha$ maps $H_\beta{}^\alpha(\mathfrak{A}^\beta)$ into a subgroup of U—therefore into the zero. As this contradicts the preceding statement, the assumption $H^\alpha(\mathfrak{A}) \subset \bar{\mathfrak{A}}^\alpha$ is false. So there is an element $\mathfrak{a}$ of $\mathfrak{A}$ whose coördinate $\mathfrak{a}^\alpha$ is not in $\bar{\mathfrak{A}}^\alpha$. As $\mathfrak{b}(\mathfrak{a}) = (\mathfrak{a}^\alpha, \mathfrak{b}^\alpha) \neq 0$, we have proved that the annihilator of $\mathfrak{A}$ in $\mathfrak{B}$ is the zero. The results of the last two paragraphs are sufficient to prove that each of the groups $\mathfrak{A}$, $\mathfrak{B}$ is the group of continuous characters of the other ([13], Lemma 5).

7. The representations of bicompact groups. We shall establish two theorems on the representations of discrete groups by direct systems of groups having finite bases. The results of the preceding section enable us to state the two dual propositions on bicompact groups.

THEOREM 7. 1. *A discrete group is always the limit group of some direct system of groups each on a finite basis. A system may be chosen so that its homomorphisms are isomorphic transformations into subgroups.*

THEOREM 7. 2. *Two direct systems of discrete groups having finite bases are equivalent if their limit groups are isomorphic.*

Let $\mathfrak{B}$ be a discrete group; and let $\{\mathfrak{B}^\alpha\}$ be the collection of all subgroups of $\mathfrak{B}$ on a finite basis. The system S will consist of these groups. If $\mathfrak{B}^\alpha \subset \mathfrak{B}^\beta$ in $\mathfrak{B}$, we introduce in S the isomorphism $H_\alpha{}^\beta$ transforming the group $\mathfrak{B}^\alpha$ into the subgroup of $\mathfrak{B}^\beta$ with which it is identified in $\mathfrak{B}$. Since any two subgroups of $\mathfrak{B}$, each on a finite basis, are contained in a third subgroup on a finite basis, it follows that S is a direct system. The isomorphism between $\mathfrak{B}$ and the limit group of S is set up in an obvious fashion. This proves Theorem 7. 1.

Let $S_1 = \{\mathfrak{B}_1{}^\alpha\}$ be a direct system of discrete groups each on a finite basis. Let $\mathfrak{B}$ be its limit group. We prove the second theorem by showing that S_1 and the system S just constructed are equivalent.

Let $\mathfrak{B}_0{}^\alpha$ be the subgroup of $\mathfrak{B}_1{}^\alpha$ whose image in $\mathfrak{B}$ is the zero. $\mathfrak{B}_0{}^\alpha$ has a finite basis; to each such basis element there is a refinement of $\mathfrak{B}_1{}^\alpha$ in which its projection is zero. Passing to a common refinement of these groups we find a refinement $\mathfrak{B}_1{}^\beta$ of $\mathfrak{B}_1{}^\alpha$ in which the image of $\mathfrak{B}_0{}^\alpha$ is the zero. Let S'_1 be the subsystem of S_1 consisting of the same groups as S_1 but contains, for each α, only those homomorphisms $\mathfrak{B}_1{}^\alpha \to \mathfrak{B}_1{}^\beta$ of S_1 under which $\mathfrak{B}_0{}^\alpha$ is mapped into zero (we except of course the identity transformations: $\mathfrak{B}_1{}^\alpha \to \mathfrak{B}_1{}^\alpha$). Then S_1 and S'_1 are equivalent systems.

Let $\mathfrak{B}'^\alpha = \mathfrak{B}_1{}^\alpha - \mathfrak{B}_0{}^\alpha$. We shall enlarge S'_1 obtaining an equivalent system S_1'' by adding the groups $\mathfrak{B}'^\alpha$. We must add also certain homomorphisms.

First $\mathfrak{B}'^{\alpha}$ is the homomorphic image of $\mathfrak{B}_1^{\alpha}$, we include this homomorphism and then any homomorphism obtained by a combination $\mathfrak{B}_1^{\gamma} \to \mathfrak{B}_1^{\alpha} \to \mathfrak{B}'^{\alpha}$. If $\mathfrak{B}_1^{\alpha} \to \mathfrak{B}_1^{\beta}$ (in S'_1), then $\mathfrak{B}'^{\alpha}$ is isomorphic with a subgroup of $\mathfrak{B}_1^{\beta}$; we include this isomorphism. Then we include any homomorphism resulting from a combination of any above: $\mathfrak{B}'^{\alpha} \to \mathfrak{B}_1^{\beta} \to \mathfrak{B}'^{\delta}$. Then S_1'' has S_1' as a complete subsystem. The system S' consisting of all groups of the type $\mathfrak{B}'^{\alpha}$ of S_1'' and all homomorphisms of S_1'' between any two such groups, also forms a complete subsystem. The system S' has the advantage that all of its homomorphisms are really isomorphic transformations (i. e. as above, $\mathfrak{B}'^{\alpha}$ is mapped isomorphically into a subgroup of $\mathfrak{B}'^{\delta}$). This follows from the fact that no non-zero element of $\mathfrak{B}'^{\alpha}$ images into the zero of $\mathfrak{B}$. Thus each group $\mathfrak{B}'^{\alpha}$ may be identified with a subgroup of $\mathfrak{B}$; and a projection $\mathfrak{B}'^{\alpha} \to \mathfrak{B}'^{\delta}$ goes over into a relation of inclusion in $\mathfrak{B}$. It is trivial that the system S' is equivalent with the system S constructed above (except for the possible repetition of the same group, it is a complete subsystem of S). We merely use inclusion relations in $\mathfrak{B}$ to define isomorphisms interlocking the two systems.

By using the theorem of the preceding section we can state two dual theorems. The character group of a free discrete group is called a *toral group.* A toral group is always the direct sum of a number of groups each isomorphic with the group $\mathfrak{X}$ (No. 6) [17; p. 370]. The number is called the *dimension* and is equal to the rank of the free group. The character group of a discrete group on a finite basis we shall call an *elementary group.* It is always the direct sum of a finite group and a toral group of finite dimension.

THEOREM 7. 3. *To each bicompact group corresponds at least one inverse system of elementary groups of which it is the limit group. A system can be chosen so that each of its groups is covered by the image of any of its refinements.*

THEOREM 7. 4. *Two inverse systems of elementary groups are equivalent if their limit groups are continuously isomorphic.*[8]

II. Homology groups over a general group of coefficients.

8. The finite complex. K denotes a finite complex (it need not be simplicial) of dimension n. L will denote a closed subcomplex of K. The

[8] In view of recent results of van Kampen: "Almost periodic functions and compact groups," *Annals of Mathematics*, vol. 37 (1936), pp. 78-91, it can be shown that these two theorems hold in the non-abelian case if Lie group replaces elementary group. The proof of the first depends on the existence in any neighborhood of the identity of a closed invariant subgroup such that the factor group is a Lie group. The proof of the second depends on the fact that any decreasing sequence of closed subgroups of a Lie group is finite.

p-cells of $K - L$ are denoted $E_p{}^i$ $(i = 1, \cdots, \alpha_p;\ p = 0, 1, \cdots, n)$. The boundary relations of K mod L are written [9]

$$F(E_p{}^i) = \eta^i{}_{p-1,j}E^j{}_{p-1}$$

where η_p is a finite matrix of integers. We assume that $\mathfrak{G}$ is a topological abelian group.

A p-chain over $\mathfrak{G}$ of K mod L is a linear form $\mathfrak{g}_iE_p{}^i$ in the p-cells of $K - L$ with coefficients $\mathfrak{g}_i \in \mathfrak{G}$. Addition is defined by: $(\mathfrak{g}_iE_p{}^i) + (\mathfrak{g}_i'E_p{}^i) = \mathfrak{g}_i''E_p{}^i$ where $\mathfrak{g}_i'' = \mathfrak{g}_i + \mathfrak{g}_i'$. In this way the p-chains over $\mathfrak{G}$ of K mod L constitute a group denoted by $\mathfrak{K}_p(K, L, \mathfrak{G})$ (We shall omit parts of the expression $\mathfrak{K}_p(K, L, \mathfrak{G})$ whenever confusion is impossible).

We introduce a topology in $\mathfrak{K}$ as follows. Let $\{U^\alpha\}$ be a complete set of neighborhoods of zero in $\mathfrak{G}$. The set V^α of those p-chains whose coefficients lie in U^α is defined to be a neighborhood of zero in $\mathfrak{K}$. The set $\{V^\alpha\}$ is taken to be complete. The set $V^\alpha(\mathfrak{k})$ obtained by adding the chain $\mathfrak{k}$ to each element of V^α is a neighborhood of $\mathfrak{k}$; and the set $\{V^\alpha(\mathfrak{k})\}$ is taken to be complete. We proceed to verify that $\mathfrak{K}$ is a topological group. Clearly $\mathfrak{k} \in V^\alpha(\mathfrak{k})$.

If $V^\alpha(\mathfrak{k})$ and $V^\beta(\mathfrak{k})$ are two neighborhoods of $\mathfrak{k}$, let U^γ be common to U^α and U^β; then $V^\gamma(\mathfrak{k}) \subset V^\alpha(\mathfrak{k})\, V^\beta(\mathfrak{k})$.

Suppose $\mathfrak{k}' \in V^\alpha(\mathfrak{k})$. Let $\mathfrak{g}_1, \cdots, \mathfrak{g}_{\alpha_p}$ be the coefficients of $\mathfrak{k}' - \mathfrak{k}$; then $\mathfrak{g}_i \in U^\alpha$. Let U^β be such that $\mathfrak{g}' \in U^\beta$ implies $\mathfrak{g}' + \mathfrak{g}_i \in U^\alpha$. Then $V^\beta(\mathfrak{k}' - \mathfrak{k}) \subset V^\alpha$ and this implies $V^\beta(\mathfrak{k}') \subset V^\alpha(\mathfrak{k})$.

Suppose now that $\mathfrak{k}' \neq \mathfrak{k}$. Let $\mathfrak{g}_i$ be a non-zero coefficient of $\mathfrak{k}' - \mathfrak{k}$; then let U^α be chosen so that $\mathfrak{g}_i$ is not the difference of any two of its elements. Then $V^\alpha(\mathfrak{k})$ and $V^\alpha(\mathfrak{k}')$ have no common point.

Finally let $V^\alpha(\mathfrak{k}_1 - \mathfrak{k}_2)$ be given. Let U^β be such that the difference of any two of its elements lies in U^α. Then $\mathfrak{k}'_1 \in V^\beta(\mathfrak{k}_1)$ and $\mathfrak{k}'_2 \in V^\beta(\mathfrak{k}_2)$ implies $\mathfrak{k}'_1 - \mathfrak{k}'_2 \in V^\alpha(\mathfrak{k}_1 - \mathfrak{k}_2)$. Thus $\mathfrak{K}$ is a Hausdorff space satisfying the axiom α) of No. 1.

Let $\mathfrak{K}^i$ be the subgroup of $\mathfrak{K}$ of those chains whose coefficients are all zero except possibly the i-th. Then it is easy to prove that $\mathfrak{K}^i$ is bicontinuously isomorphic with $\mathfrak{G}$ and $\mathfrak{K}$ with the direct sum $\sum \mathfrak{K}^i$. Thus

LEMMA 8.1. *$\mathfrak{K}_p(K, L, \mathfrak{G})$ is the direct sum of α_p groups each bicontinuously isomorphic with $\mathfrak{G}$.*

The *boundary* mod L of a p-chain is defined by

[9] When a latin index occurs twice in a term, once above and once below, it is to be summed over its range.

$$F(\mathfrak{g}_i E_p{}^i) = \mathfrak{g}_i \eta^i{}_{p-1,j} E^j{}_{p-1}. \tag{8.1}$$

It is clear that F is a homomorphic mapping of $\mathfrak{K}_p$ into $\mathfrak{K}_{p-1}$. In fact.

LEMMA 8.2. *The boundary operator is a continuous homomorphic mapping of $\mathfrak{K}_p$ into $\mathfrak{K}_{p-1}$.*

Let V^α be a neighborhood of zero in $\mathfrak{K}_{p-1}$. Now $\mathfrak{g}_i \eta^i{}_{p-1,j}$ is a continuous function in the $\mathfrak{g}_i$ simultaneously and is zero for each $\mathfrak{g}_i = 0$. Therefore there is a neighborhood $U_j{}^\beta$ of zero in $\mathfrak{G}$ such that $\mathfrak{g}_i \eta^i{}_{p-1,j} \in U^\alpha$ whenever each $\mathfrak{g}_i \in U_j{}^\beta$. Let $U^\beta \subset \prod_j U_j{}^\beta$. Then the neighborhood V^β of zero in $\mathfrak{K}_p$ determined by U^β is such that $F(V^\beta) \subset V^\alpha$. Thus F is continuous at the zero of $\mathfrak{K}_p$. As F is homomorphic, it is continuous everywhere.

It follows from the lemma that the subgroup of $\mathfrak{K}_p$ mapped by F into the zero of $\mathfrak{K}_{p-1}$ is closed in $\mathfrak{K}_p$. This group, $\mathfrak{C}_p(K, L, \mathfrak{G})$, is called the *group of p-cycles over* $\mathfrak{G}$ of K mod L. The image under F of $\mathfrak{K}_{p+1}$ in $\mathfrak{K}_p$ is a subgroup, $\mathfrak{B}_p(K, L, \mathfrak{G})$, of $\mathfrak{K}_p$ called the *group of bounding p-cycles over* $\mathfrak{G}$ of K mod L. Since $\eta_p \eta_{p-1} = 0$, we have $\mathfrak{B}_p \subset \mathfrak{C}_p$.

DEFINITION 1. The *p-th homology group over* $\mathfrak{G}$ of K mod L is

$$\mathfrak{H}_p(K, L, \mathfrak{G}) = \mathfrak{C}_p(K, L, \mathfrak{G}) - \bar{\mathfrak{B}}_p(K, L, \mathfrak{G})$$

where $\bar{\mathfrak{B}}$ is the closure of $\mathfrak{B}$.

We will see shortly the conditions that $\mathfrak{G}$ must satisfy in order that $\mathfrak{B}$ should be closed. We proceed to give the structure of $\mathfrak{H}$ in terms of $\mathfrak{G}$ and the Betti numbers and torsion coefficients of K mod L.

Let λ_p $(p = 0, 1, \cdots, n)$ be a unimodular matrix of integers of order α_p, and let $\bar{\lambda}_p$ denote its inverse: $\lambda_p \bar{\lambda}_p = 1$. Let $G_p{}^j$ denote the linear form $\lambda_{pi}{}^j E_p{}^i$. We define for each G a boundary

$$F(G_p{}^i) = \zeta^i{}_{p-1,j} G^j{}_{p-1}, \qquad \zeta_{p-1} = \lambda_p \eta_{p-1} \bar{\lambda}_{p-1}. \tag{8.2}$$

Let $\mathfrak{K}'_p$ denote the group of linear forms in the $G_p{}^i$ with coefficients from $\mathfrak{G}$. We define a topology for $\mathfrak{K}'_p$ as we defined one for $\mathfrak{K}_p$. In analogy with (8.1) the transformation

$$F(\mathfrak{g}_i G_p{}^i) = \mathfrak{g}_i \zeta^i{}_{p-1,j} G^j{}_{p-1} \tag{8.3}$$

is a continuous homomorphism of $\mathfrak{K}'_p$ into $\mathfrak{K}'_{p-1}$. The transformation

$$f(\mathfrak{g}_i G_p{}^i) = \mathfrak{g}_i \lambda_{pj}{}^i E_p{}^j$$

is a homomorphism of $\mathfrak{K}'_p$ into $\mathfrak{K}_p$ which is continuous (compare with proof of Lemma 8.2). It has the continuous inverse

$$f'(\mathfrak{g}_j E_p{}^j) = \mathfrak{g}_j \bar{\lambda}_{pk}{}^j G_p{}^k,$$

and is therefore a bicontinuous isomorphism between $\mathfrak{K}_p$ and $\mathfrak{K}'_p$. Comparing (8.2) with (8.1) we see that the boundary operator F commutes with this isomorphism. Therefore the groups $\mathfrak{K}'_p$ and the boundary operator (8.3) may be used to define the homology groups over $\mathfrak{G}$ of K mod L.

It is known [15; p. 27] that there exist unimodular matrices λ_p such that the matrices ζ_p of (8.2) are all in quasi-canonical form (i. e. at most one non-zero element in any row or column). In this case we can distinguish five types of G's (which are given the new notation $a, b, \cdots$ respectively) so that the boundary relations (8.2) assume the form

$$(8.4)\quad \begin{array}{lll} F(e^h{}_{p+1}) = a_p{}^h, & F(a_p{}^h) = 0, & (h = 1, \cdots, \omega_p), \\ F(d^i{}_{p+1}) = \theta_p{}^i b_p{}^i, & F(b_p{}^i) = 0, & (i = 1, \cdots, \tau_p), \\ & F(c_p{}^j) = 0, & (j = 1, \cdots, R_p), \\ & F(d_p{}^k) = \theta^k{}_{p-1} b^k{}_{p-1}, & (k = 1, \cdots, \tau_{p-1}), \\ & F(e_p{}^l) = a^l{}_{p-1}, & (l = 1, \cdots, \omega_{p-1}), \end{array}$$

where the $\theta_p{}^i$ $(i = 1, \cdots, \tau_p)$ are the invariant factors of η_p different from 1, $R_p = \alpha_p - \rho_p - \rho_{p-1}$ $(\rho_p = \text{rank } \eta_p)$, and $\omega_p = \rho_p - \tau_p$. If $\mathfrak{A}_p{}^h$ denotes the subgroup of $\mathfrak{K}_p(\mathfrak{G})$ of elements of the form $\mathfrak{g} a_p{}^h$, etc., then

$$\mathfrak{K}_p(K, L, \mathfrak{G}) = \sum \mathfrak{A}_p{}^h + \sum \mathfrak{B}_p{}^i + \sum \mathfrak{C}_p{}^j + \sum \mathfrak{D}_p{}^k + \sum \mathfrak{E}_p{}^l.$$

If $\mathfrak{D}'_p{}^k$ is the subgroup of $\mathfrak{D}_p{}^k$ of elements of orders dividing $\theta^k{}_{p-1}$, then

$$\mathfrak{C}_p(K, L, \mathfrak{G}) = \sum \mathfrak{A}_p{}^h + \sum \mathfrak{B}_p{}^i + \sum \mathfrak{C}_p{}^j + \sum \mathfrak{D}'_p{}^k.$$

If $\mathfrak{B}'_p{}^i$ is the subgroup of elements of $\mathfrak{B}_p{}^i$ divisible by $\theta_p{}^i$ and $\bar{\mathfrak{B}}'_p{}^i$ is its closure, then

$$\bar{\mathfrak{B}}_p(K, L, \mathfrak{G}) = \sum \mathfrak{A}_p{}^h + \sum \bar{\mathfrak{B}}'_p{}^i.$$

Referring to Lemma 1.2, we can state

THEOREM 8. *If R_p, $\theta^i{}_{p-1}$ $(i = 1, \cdots, \tau_{p-1};\ p = 0, 1, \cdots, n)$ are the Betti numbers and torsion coefficients of the finite n-complex K mod L, then $\mathfrak{H}_p(K, L, \mathfrak{G})$ can be continuously decomposed into the direct sum*

$$(8.5)\qquad \sum_{i=1}^{\tau_p} \mathfrak{G}^*{}_p{}^i + \sum_{j=1}^{R_p} \mathfrak{G}_p{}^j + \sum_{k=1}^{\tau_{p-1}} \mathfrak{G}'_p{}^k$$

where $\mathfrak{G}^*{}_p{}^i$ is the group obtained by reducing $\mathfrak{G}$ modulo the closure of the subgroup of elements divisible by $\theta_p{}^i$, each $\mathfrak{G}_p{}^j$ is bicontinuously isomorphic with $\mathfrak{G}$, and $\mathfrak{G}'{}_p{}^k$ is the subgroup of $\mathfrak{G}$ of elements whose orders divide $\theta^k{}_{p-1}$.

The condition that $\mathfrak{B}_p$ should be closed is clearly that each $\mathfrak{B}'{}_p{}^i$ should be closed. Thus the following requirement on $\mathfrak{G}$ is sufficient for $\mathfrak{B}_p$ to be closed:

DIVISION-CLOSURE PROPERTY. For each positive integer m, the subgroup of $\mathfrak{G}$ of elements divisible by m is closed.

In general this condition is also necessary; for, if $m\mathfrak{G}$ is not closed for some integer m, the complex K composed of a circle bounding a 2-cell where sets of m equally spaced points on the circle are identified has the property that $\mathfrak{B}_1(K, \mathfrak{G})$ is not closed.

If $\mathfrak{G}$ has the division-closure property, then the algebraic structure of $\mathfrak{H}_p(\mathfrak{G})$ depends only on the algebraic structure of $\mathfrak{G}$. Thus, if $\mathfrak{G}$ and $\mathfrak{G}'$ are isomorphic (though not continuously so) and both have the division-closure property, then $\mathfrak{H}_p(\mathfrak{G})$ and $\mathfrak{H}_p(\mathfrak{G}')$ are isomorphic.

It is easy to prove that discrete groups, bicompact groups, and vector groups have the division-closure property. Likewise the groups of rational numbers. If two groups have the property, so also does their direct sum. A connected, locally-bicompact group has the property. On the other hand, the group of rational numbers of the form $p/2^n$ (p and n are integers) with the topology it has as a subset of the linear continuum does not have the division-closure property. For 1/2 is not divisible (in the group) by 3, while the elements divisible by 3 are everywhere dense.

We state some simple consequences of Theorem 8.

COROLLARY 8.1. *If $\mathfrak{G}$ is bicompact, so also is $\mathfrak{H}_p(\mathfrak{G})$. If $\mathfrak{G}$ is an elementary group (No. 7), so is $\mathfrak{H}_p(\mathfrak{G})$. If $\mathfrak{G}$ is the group $\mathfrak{X}$ of real numbers reduced mod 1, then $\mathfrak{H}_p(\mathfrak{G})$ is the direct sum of a toral group of dimension R_p and of a finite group isomorphic with the torsion group over the integers of one less dimension.*

Thus the Betti numbers and torsion coefficients are invariants of the mod 1 homology groups as well as the integer homology groups; hence

COROLLARY 8.2. *The group $\mathfrak{X}$ and the group of integers are each universal coefficient groups for the homology theory of a finite complex.*

DEFINITION 2. $\mathfrak{T}_p(K, L, \mathfrak{G}) = \sum \mathfrak{G}^*{}_p{}^i$ is called the *torsion group over* $\mathfrak{G}$. $\mathfrak{S}_p(K, L, \mathfrak{G}) = \mathfrak{H}_p(\mathfrak{G}) - \mathfrak{T}_p(\mathfrak{G})$ is called the *reduced homology group.* $\sum \mathfrak{G}'{}_p{}^k$ is called the *group of impure cycles* (see Čech [9]).

REMARK. The decomposition (8.5) in general cannot be specified in terms of invariants of $\mathfrak{H}_p(\mathfrak{G})$. For example there may exist a bicontinuous isomorphism of $\mathfrak{H}_p(\mathfrak{G})$ into itself which carries the torsion group into another subgroup. The decomposition is imposed by the structure of the complex, not by that of the group.

9. The abstract space. We describe briefly the procedure of Čech [8]. Let A be a set of points. A finite covering ϕ of A is a finite collection of subsets of A whose sum is A. A finite covering is a refinement of another if every set of the first is contained in some set of the second. In particular, any finite covering is a refinement of itself.

Each finite covering may be regarded as a complex K (the nerve of the covering): the sets are the vertices, a collection of $n+1$ vertices are the vertices of an n-simplex if the corresponding sets have a non-vacuous intersection. The intersection is associated with the n-simplex. If B is a subset of A, it determines a closed subcomplex L of K where a simplex is in L if and only if its associated set meets B.

If ϕ^2 is a refinement of ϕ^1 $(\phi^2 \to \phi^1)$ and if K^2, L^2 and K^1, L^1 are the corresponding complexes and subcomplexes, we may define a simplicial mapping π of K^2 into K^1. To each set of ϕ^2 we associate a definite one of the sets of ϕ^1 which contains it; then π maps the vertex of K^2 corresponding to the set into the vertex if K^1 corresponding to the associated containing set. It is not difficult to see that π is a simplicial mapping and carries L^2 into L^1. The projection π obtained in this manner is called an allowable projection. In general there are many allowable projections.

If π maps a simplex σ_p of $K^2 - L^2$ into a simplex of L^1 or a simplex of dimension $< p$ of $K^1 - L^1$ we write algebraically $\pi(\sigma_p) = 0$. If σ_p is carried into a simplex τ_p of $K^1 - L^1$ we write $\pi(\sigma_p) = \tau_p$. Then, for any chain $\mathfrak{g}_i\sigma_p{}^i$ of $K^2 - L^2$ we define $\pi(\mathfrak{g}_i\sigma_p{}^i) = \mathfrak{g}_i\pi(\sigma_p{}^i)$. In this way we define a homomorphism π of $\mathfrak{K}_p(K^2, L^2, \mathfrak{G})$ into $\mathfrak{K}_p(K^1, L^1, \mathfrak{G})$. Since π (like F) is determined by a finite matrix of integers, we can prove that π is continuous (just as we proved that F is continuous). Since $\pi F(\sigma_p) = F\pi(\sigma_p)$ we find that π maps cycles (bounding cycles) into cycles (bounding cycles). Thus π induces a continuous homomorphism H of $\mathfrak{H}_p(K^2, L^2, \mathfrak{G})$ into $\mathfrak{H}_p(K^1, L^1, \mathfrak{G})$.

If π' is another allowable projection of K^2 into K^1, and c_p is a cycle of K^2 mod L^2, then, as shown by Čech, $\pi(c_p) \sim \pi'(c_p)$ mod L^1. Thus any allowable projection induces the same homomorphism H.

Now let Φ be a family $\{\phi^\alpha\}$ of finite coverings of A satisfying the condition that any two members of Φ have a common refinement in Φ. If $\phi^\beta \to \phi^\alpha$, let $H_\beta{}^\alpha$ be the induced homomorphism of $\mathfrak{H}_p{}^\beta = \mathfrak{H}_p(K^\beta, L^\beta, \mathfrak{G})$ into $\mathfrak{H}_p{}^\alpha$. The

limit group $\mathfrak{H}_p$ of the inverse system $\{\mathfrak{H}_p{}^\alpha\}$ is called the *p-th homology over* $\mathfrak{G}$ of the family Φ mod B.

If A is a topological space and Φ is the family of all finite coverings of A by closed sets, $\mathfrak{H}_p$ is called the *p-th homology group over* $\mathfrak{G}$ of A mod B, written $\mathfrak{H}_p(A, B, \mathfrak{G})$. By definition $\mathfrak{H}_p$ is a topological invariant of A (i. e. homeomorphic spaces have bicontinuously isomorphic homology groups).

REMARK I. We may also choose Φ to be the family of all finite coverings of A by open sets, and $\mathfrak{H}$ is once again a topological invariant. Čech [8; p. 180] has proved that, in a normal space, the two homology theories are equivalent.[10] It must be remarked that the proof of the theorem of the next section uses closed coverings in an essential way. Thus we have no proof of the analogous theorem for the homology group of a non-normal space based on open coverings. This remark applies also to Theorem 12.

II. If $\mathfrak{G}$ is bicompact, by Theorem 2.1, $\mathfrak{H}_p(A, B, \mathfrak{G})$ is bicompact. According to Čech, a cycle of the covering ϕ^α is *essential* if it is homologous to the projection of a cycle of ϕ^β for every refinement ϕ^β of ϕ^α. Čech proved for rational coefficients the fundamental existence theorem that an essential cycle of ϕ^α is a representative of the α-th coördinate of some element of $\mathfrak{H}_p$. In Lemma 2.1, we established this fundamental existence theorem for bicompact coefficient groups. According to Čech, a refinement ϕ^β of ϕ^α is *normal relative to* $\mathfrak{G}$ if the projection in ϕ^α of each p-cycle over $\mathfrak{G}$ of ϕ^β is essential. By Corollary 8.1, $\mathfrak{H}_p(K^\alpha, L^\alpha, \mathfrak{G})$ is an elementary group if $\mathfrak{G}$ is an elementary group. In proving Theorem 7.4, we showed that any elementary group in an inverse system of bicompact groups has a normal refinement. Thus normal refinements always exist relative to coefficient groups which are elementary. If $\mathfrak{G}$ is bicompact but not elementary, then in general normal refinements do not exist. However Lemma 2.2 states that, for a bicompact $\mathfrak{G}$, we may obtain refinements as close to being normal as we please.

III. If K is a finite complex and L a closed subcomplex, we have two definitions of $\mathfrak{H}_p(K, L, \mathfrak{G})$, the first, considering K as a complex, the second, considering K as a topological space. One establishes a bicontinuous isomorphism between these two groups as follows. Let $\{K^m\}$ be a sequence of barycentric subdivisions of K. The set ϕ^m of the stars of vertices of K^m form a finite covering of K by open sets whose associated complex is K^m. The family $\Phi = \{\Phi^m\}$ is a complete family in the sense that every covering of K by open sets has a refinement in Φ (see Remark I, K is normal). It is then proved in the usual way that the projection of K^{m+1} into K^m induces an iso-

[10] Čech's proof assumes that A is completely normal, but this is not necessary.

morphism of $\mathfrak{H}_p(K^{m+1}, L^{m+1}, \mathfrak{G})$ into the whole of $\mathfrak{H}_p(K^m, L^m, \mathfrak{G})$. It is not difficult to see that this isomorphism is bicontinuous. Then the limit group is bicontinuously isomorphic with each group of the sequence.

10. General direct sum theorem. The *torsion group over* $\mathfrak{G}$ of A mod B is the subgroup of $\mathfrak{H}$ of those elements represented in each covering complex by elements of the torsion group of that complex (Def. 8.2). We shall prove [11]

THEOREM 10.1. *$\mathfrak{H}_p(A, B, \mathfrak{G})$ is the direct sum of the torsion group $\mathfrak{T}_p$ and a reduced homology group $\mathfrak{S}_p$.*

In order to exhibit this decomposition of $\mathfrak{H}$ we shall define a canonical inverse system of groups of chains which carries the homology theory of A mod B. This canonical system will also be used in the next section.

To obtain an inverse system of chains we use a type of closed covering introduced by Kurosh [14]. We then consider the direct system of groups of dual chains (in the sense of Alexander [1]) determined by these coverings. To avoid a digression into the theory of dual chains, we define this direct system in terms of the inverse system in an invariant fashion using a type of argument due to Pontrjagin [16]. In the limit group of the direct system, a decomposition into a direct sum is established. Herein is contained a direct sum theorem for the dual homology group. We then argue backwards and define in each group of the inverse system a canonical basis so that the matrices exhibiting the homomorphisms of the system have a simple form. Finally we interpolate certain auxiliary groups of chains into the system. The subsystem of auxiliary groups is the desired canonical system and the direct sum theorem is immediate.

In the manner of Kurosh [14], we consider finite coverings of A by closed sets having the property that each set is the closure of its interior and the interiors of distinct sets are non-overlapping. Let Φ' be the aggregate of such coverings of A. It is easy to prove that any finite covering of A by closed sets has a refinement in Φ'. Thus Φ' is a complete subfamily of Φ, and we may restrict our considerations to Φ'. These coverings of A have the following advantages:

1°. If $\phi^\alpha \rightarrow \phi^\beta$, each set of ϕ^β is an exact sum of sets of ϕ^α.

2°. If $\phi^\alpha \rightarrow \phi^\beta$, each set of ϕ^α is contained in only one set of ϕ^β.

---

[11] For the special case of a compact metric space and homology groups based on integer coefficients, this theorem was proved by Alexander and Cohen [2].

From 1° we deduce that a covering is a refinement of only finitely many others. Thus a covering possesses an immediate refinement in the sense that there are no coverings of Φ' between the two. In general a covering possesses many immediate refinements. From 2° it follows that the projection of the nerves $K^\alpha \to K^\beta$ is uniquely determined. As a consequence the groups of chains $\mathfrak{K}_p{}^\alpha = \mathfrak{K}_p(K^\alpha, L^\alpha, \mathfrak{J})$ ($\mathfrak{J}$ = group of integers) of the covering complexes K^α form an inverse system. From 1° and 2° it is proved that K^β is completely covered by the image of K^α. Then the homomorphism of $\mathfrak{K}_p{}^\alpha$ into $\mathfrak{K}_p{}^\beta$ is determined by a matrix reducible to the form $\left\| \begin{matrix} 1 \\ 0 \end{matrix} \right\|$.

Let $\mathfrak{K}^{*\alpha}$ be the group of homomorphic mappings [12] of $\mathfrak{K}^\alpha$ into $\mathfrak{J}$. $\mathfrak{K}^{*\alpha}$ is isomorphic to $\mathfrak{K}^\alpha$, and $\mathfrak{K}^\alpha$ is the group of homomorphic mappings of $\mathfrak{K}^{*\alpha}$ into $\mathfrak{J}$. The element of $\mathfrak{J}$ determined by an element $\mathfrak{k} \in \mathfrak{K}^\alpha$ and $\mathfrak{k}^* \in \mathfrak{K}^{*\alpha}$ is denoted $(\mathfrak{k}, \mathfrak{k}^*)$. Following Alexander [1], we refer to the elements of $\mathfrak{K}^{*\alpha}$ as dual chains. An independent basis $\mathfrak{k}^1, \mathfrak{k}^2, \cdots, \mathfrak{k}^h$ in $\mathfrak{K}^\alpha$ determines a dual basis in $\mathfrak{K}^{*\alpha}$ (and conversely) as follows: $\mathfrak{k}^{*i}$ $(i = 1, \cdots, h)$ is defined by $(\mathfrak{k}^j, \mathfrak{k}^{*i}) = \delta_j{}^i$ $(\delta_i{}^i = 1,\ \delta_j{}^i = 0\ (i \neq j))$.

The boundary homomorphism F of $\mathfrak{K}^\alpha{}_{p+1}$ into $\mathfrak{K}_p{}^\alpha$ determines a homomorphism F^* of $\mathfrak{K}^*{}_p{}^\alpha$ into $\mathfrak{K}^{*\alpha}_{p+1}$ as follows. If $\mathfrak{k}^*{}_p \in \mathfrak{K}^*{}_p{}^\alpha$, the function $\mathfrak{k}^*{}_p(\mathfrak{k}_{p+1}) = (F(\mathfrak{k}_{p+1}), \mathfrak{k}^*{}_p)$ is a homomorphism of $\mathfrak{K}^\alpha{}_{p+1}$ into $\mathfrak{J}$; it can therefore be identified with an element $\mathfrak{k}^*{}_{p+1} \in \mathfrak{K}^{*\alpha}_{p+1}$. Then $F^*(\mathfrak{k}^*{}_p)$ is the element $\mathfrak{k}^*{}_{p+1}$ so obtained. By definition

$$(10.1) \qquad (F(\mathfrak{k}_{p+1}), \mathfrak{k}^*{}_p) = (\mathfrak{k}_{p+1}, F^*(\mathfrak{k}^*{}_p))$$

for arbitrary $\mathfrak{k}_{p+1} \in \mathfrak{K}^\alpha{}_{p+1}$ and $\mathfrak{k}^*{}_p \in \mathfrak{K}^*{}_p{}^\alpha$. If F is given by a matrix η in terms of fixed bases in $\mathfrak{K}^\alpha{}_{p+1}$ and $\mathfrak{K}^\alpha{}_p$, and F^* by a matrix ζ in terms of the dual bases, by (10.1)

$$(F(\mathfrak{k}^i{}_{p+1}), \mathfrak{k}^*{}_p{}^j) = \eta_k{}^i(\mathfrak{k}_p{}^k, \mathfrak{k}^*{}_p{}^j) = \eta_j{}^i = (\mathfrak{k}^i{}_{p+1}, F^*(\mathfrak{k}^*{}_p{}^j)) = \zeta_k{}^j(\mathfrak{k}^i{}_{p+1}, \mathfrak{k}^{*k}_{p+1}) = \zeta_i{}^j.$$

Therefore ζ is the transpose of η.

If $\mathfrak{K}^\beta \to \mathfrak{K}^\alpha$, this homomorphism H defines in a similar way a dual H^* transforming $\mathfrak{K}^{*\alpha}$ into $\mathfrak{K}^{*\beta}$ satisfying

$$(10.2) \qquad (H(\mathfrak{k}^\beta), \mathfrak{k}^{*\alpha}) = (\mathfrak{k}^\beta, H^*(\mathfrak{k}^{*\alpha})).$$

In this way $\{\mathfrak{K}^{*\alpha}\}$ becomes a direct system. From the relation $FH = HF$ and the relations (10.1) and (10.2) we can deduce the dual relation

$$(10.3) \qquad F^*H^* = H^*F^*.$$

[12] The use of this notion is due to Pontrjagin [16]. Proofs of the statements of this paragraph can be found in his paper.

Since the matrix of H^* is the transpose of that of H, by a remark above, bases can be chosen in $\mathfrak{K}^{*\alpha}$ and $\mathfrak{K}^{*\beta}$ so that this matrix has the form $\| 1, 0 \|$. Therefore H^* maps $\mathfrak{K}^{*\alpha}$ isomorphically into a direct summand of $\mathfrak{K}^{*\beta}$. We shall use this to prove that the limit group $\mathfrak{K}^*$ of $\{\mathfrak{K}^{*\alpha}\}$ (No. 5) is a free group. We construct a free basis. Choose a free basis in $\mathfrak{K}^{*1}$ and form its image in $\mathfrak{K}^*$. These elements are independent in $\mathfrak{K}^*$; for the contrary would imply a refinement $\mathfrak{K}^{*\alpha}$ of $\mathfrak{K}^{*1}$ into which $\mathfrak{K}^{*1}$ is not isomorphically mapped. Suppose we have found an independent set S_β in $\mathfrak{K}^*$ generating a subgroup $\mathfrak{K}^*_\beta$ which contains the image of each $\mathfrak{K}^{*\alpha}$ for $\alpha < \beta$. Let $\mathfrak{K}^*{}_0{}^\beta$ be the subgroup of $\mathfrak{K}^{*\beta}$ imaging into $\mathfrak{K}^*_\beta$. It follows that there exist groups $\mathfrak{K}^{*\alpha_i}$ $(\alpha_i < \beta,\ i = 1, \cdots, k)$ and a refinement $\mathfrak{K}^{*\gamma}$ of $\mathfrak{K}^{*\beta}$ and the $\mathfrak{K}^{*\alpha_i}$ in which the images of $\mathfrak{K}^{*\beta}$ and $\mathfrak{K}^{*\alpha_i}$ intersect in the image of $\mathfrak{K}^*{}_0{}^\beta$. Since these transformations are isomorphisms into direct summands of $\mathfrak{K}^{*\gamma}$, it follows that $\mathfrak{K}^*{}_0{}^\beta$ is a direct summand of $\mathfrak{K}^{*\beta}$. We can therefore find a set of independent elements of $\mathfrak{K}^{*\beta}$ not in $\mathfrak{K}^*{}_0{}^\beta$ which with the latter group generate $\mathfrak{K}^{*\beta}$. We form the image in $\mathfrak{K}^*$ of this set and adjoin it to S_β to obtain $S_{\beta+1}$. The remaining argument is obvious. It is also clear that $\mathfrak{K}^{*\alpha}$ images into a direct summand of $\mathfrak{K}^*$.

From the relation (10. 3) it is seen that the system of homomorphisms $\mathfrak{K}^*{}_p{}^\alpha \to \mathfrak{K}^{*\alpha}_{p+1}$ defines a homomorphism F^* of $\mathfrak{K}^*_p$ into $\mathfrak{K}^*_{p+1}$ (if $\{\mathfrak{k}^*{}_p{}^\alpha\} \in \mathfrak{K}^*_p$, $F^*(\{\mathfrak{k}^*{}_p{}^\alpha\}) = \{F^*(\mathfrak{k}^*{}_p{}^\alpha)\} \in \mathfrak{K}^*_{p+1}$). As $\mathfrak{K}^*_{p+1}$ is a free group, $\mathfrak{K}^*_p$ decomposes into a direct sum $\mathfrak{C}^*_p + \mathfrak{D}^*_p$ such that $F^*(\mathfrak{C}^*_p) = 0$ and F^* maps $\mathfrak{D}^*_p$ isomorphically into a subgroup of $\mathfrak{K}^*_{p+1}$. We digress for the moment to state a theorem on the dual homology group.

If $\mathfrak{G}$ is an arbitrary abelian group (no topology assumed) the group $\mathfrak{K}^*_p(\mathfrak{G})$ of *dual chains over* $\mathfrak{G}$ of A mod B is the set of all finite linear forms in elements of $\mathfrak{K}^*_p$ with coefficients in $\mathfrak{G}$. The *boundary* of a dual chain over $\mathfrak{G}$ is defined $F^*(\mathfrak{g}_i\mathfrak{k}^{*i}) = \mathfrak{g}_i F(\mathfrak{k}^{*i})$. It is obvious that $\mathfrak{C}^*_p(\mathfrak{G})$ consists entirely of dual cycles; these are called the *pure cycles*. Certain elements of $\mathfrak{D}^*_p(\mathfrak{G})$ may be cycles; this subgroup is called the *group of impure cycles*. Further, since $F^*F^* = 0$, $\mathfrak{C}^*_p(\mathfrak{G})$ contains the subgroup of bounding cycles. As the *dual homology group* is obtained by reducing the group of cycles modulo the bounding cycles, we obtain

THEOREM 10. 2. *The dual homology group over* $\mathfrak{G}$ *of* A *mod* B *is expressible as a direct sum of a group of impure cycles and the group of those elements representable as linear forms with coefficients from* $\mathfrak{G}$ *in dual cycles with integer coefficients.*[13]

---

[13] This theorem together with an argument of Čech [9] provides a direct proof that

We return to the former discussion. Let us choose an independent basis for $\mathfrak{D}^*_p$. The group $\mathfrak{K}^*{}_p^\alpha$ decomposes into a direct sum $\mathfrak{C}^*{}_p^\alpha + \mathfrak{D}^*{}_p^\alpha$ where $F^*(\mathfrak{C}^*{}_p^\alpha) = 0$ and this does not hold for any non-zero element of $\mathfrak{D}^*{}_p^\alpha$. The latter group may be chosen in many ways. We make the following selection. Choose those elements of $\mathfrak{K}^*{}_p^\alpha$ which image into basis elements of $\mathfrak{D}^*_p$. These we call *permanent* basis elements of $\mathfrak{D}^*{}_p^\alpha$. They generate a direct summand of $\mathfrak{K}^*{}_p^\alpha$ which does not meet $\mathfrak{C}^*{}_p^\alpha$. We may therefore adjoin further *temporary* basis elements so as to obtain a complete independent basis for $\mathfrak{K}^*{}_p^\alpha$.

We now pass to a complete subsystem of $\{\mathfrak{K}^*{}_p^\alpha\}$. The subsystem will contain all the groups of the system but will contain only a subclass of the homomorphisms. If $\mathfrak{K}^*{}_p^\alpha \rightarrow \mathfrak{K}^*{}_p^\beta$, this homomorphism is included in the subsystem only if the image in $\mathfrak{K}^*{}_p^\beta$ of an arbitrary element of $\mathfrak{K}^*{}_p^\alpha$ may be reduced to a cycle by the addition of a linear form in permanent basis element of $\mathfrak{D}^*{}_p^\beta$.

We prove the existence of such a refinement $\mathfrak{K}^*{}_p^\beta$ of $\mathfrak{K}^*{}_p^\alpha$ as follows. Let $\bar{\mathfrak{D}}^*{}_p^\alpha$ be the subgroup of $\mathfrak{D}^*{}_p^\alpha$ imaging into $\mathfrak{C}^*_p$. As $\bar{\mathfrak{D}}^*{}_p^\alpha$ has a finite basis, in some refinement $\mathfrak{K}^*{}_p^\gamma$ of $\mathfrak{K}^*{}_p^\alpha$ the image of $\bar{\mathfrak{D}}^*{}_p^\alpha$ lies in $\mathfrak{C}^*{}_p^\gamma$. The image of each element of $\mathfrak{K}^*{}_p^\alpha$ in $\mathfrak{K}^*_p$ may be reduced to an element of $\mathfrak{C}^*_p$ by adding a finite form in basis elements of $\mathfrak{D}^*_p$. Since $\mathfrak{K}^*{}_p^\alpha$ has a finite basis, a finite number of the basis elements of $\mathfrak{D}^*_p$ will suffice for all elements of $\mathfrak{K}^*{}_p^\alpha$. There will exist therefore a $\mathfrak{K}^*{}_p^\delta$ whose image in $\mathfrak{K}^*_p$ contains these basis elements of $\mathfrak{D}^*_p$. Then a common refinement $\mathfrak{K}^*{}_p^\beta$ of $\mathfrak{K}^*{}_p^\gamma$ and $\mathfrak{K}^*{}_p^\delta$ obviously satisfies the above condition. To see that any two groups have a common refinement relative to the subsystem, we need only choose for each a proper refinement in the subsystem then a refinement of the latter two in the original system.

Let us perform this operation for $p = 0$. The subsystem obtained has a corresponding complete subsystem in $\{\mathfrak{K}^*{}_1^\alpha\}$. In this subsystem we may perform the same operation. The new subsystem has corresponding subsystems of dimensions 0 and 2. In the latter, the same operation can be performed. In general, if $n \geqq 0$ is an integer, systems $\{\mathfrak{K}_p^\alpha\}$ can be determined for $p \leqq n$ satisfying the above condition. Future considerations will be confined to these subsystems.

For each α, let us choose an independent basis for $\mathfrak{C}^{*\alpha}$. Suppose $\mathfrak{K}^{*\alpha} \rightarrow \mathfrak{K}^{*\beta}$; due to the conditions satisfied by $\mathfrak{K}^{*\beta}$, this homomorphism expressed in terms of the bases which have been defined in these groups has the matrix form

---

the integers form a universal coefficient group for the dual homology theory. However this result can also be obtained from our Theorem 12 by passing to character groups.

$$\begin{array}{cc|cc|l}
 & \mathfrak{C}^{*\beta} & \mathfrak{D}^{*\beta} & & \\
\mathfrak{C}^{*\alpha} & W' & 0 & 0 & \\
\hline
\mathfrak{D}^{*\alpha} & 0 & X' & 0 & \text{perm.} \\
 & Y' & Z' & 0 & \text{temp.} \\
 & & \text{perm.} & \text{temp.} &
\end{array} \qquad X' = \| 1, 0 \|.$$

In $\mathfrak{K}^{\alpha}$, for each α, let us choose the basis dual to the basis of $\mathfrak{K}^{*\alpha}$. Then the homomorphism $\mathfrak{K}^{\beta} \rightarrow \mathfrak{K}^{\alpha}$ assumes the form

$$\tag{10.4} \begin{array}{cc|cc|l}
 & \mathfrak{C}^{\alpha} & \mathfrak{B}^{\alpha} & & \\
\mathfrak{C}^{\beta} & W & 0 & Y & \\
\hline
\mathfrak{B}^{\beta} & 0 & X & Z & \text{perm.} \\
 & 0 & 0 & 0 & \text{temp.} \\
 & & \text{perm.} & \text{temp.} &
\end{array}$$

where the basis of $\mathfrak{C}^{\alpha}$ is dual to that of $\mathfrak{C}^{*\alpha}$, and the basis of $\mathfrak{B}^{\alpha}$ is dual to that of $\mathfrak{D}^{*\alpha}$ and has been divided into two parts according to whether or not an element is dual to a permanent or temporary element.

The group $\mathfrak{B}_p{}^{\alpha}$ is the group of cycles which bound or have bounding multiples. To see this, let ρ be the rank of $\mathfrak{D}^*{}_p{}^{\alpha}$ ($=$ rank of $\mathfrak{B}_p{}^{\alpha}$). As F^* maps $\mathfrak{D}^*{}_p{}^{\alpha}$ isomorphically into $\mathfrak{K}^{*\alpha}_{p+1}$ the smallest direct summand $\mathfrak{B}^{*\alpha}_{p+1}$ of the latter group containing its image has the rank ρ. We choose an independent basis of $\mathfrak{K}^{*\alpha}_{p+1}$ containing a basis of $\mathfrak{B}^{*\alpha}_{p+1}$, and then we form the dual basis in $\mathfrak{K}^{\alpha}{}_{p+1}$. The group $\mathfrak{D}^{\alpha}{}_{p+1}$ generated by the chains dual to basis elements of $\mathfrak{B}^{*\alpha}_{p+1}$ has the rank ρ. Since F and F^* are dual, F maps $\mathfrak{D}^{\alpha}{}_{p+1}$ isomorphically into a subgroup of $\mathfrak{B}_p{}^{\alpha}$. As $\mathfrak{B}_p{}^{\alpha}$ has the rank ρ, the statement is proved.

Suppose $\mathfrak{K}^{\beta}$ is a proper refinement of $\mathfrak{K}^{\alpha}$. We shall introduce between these two groups a third auxiliary group $\mathfrak{K}^{\beta\alpha}$. It shall be the direct sum $\mathfrak{C}^{\beta\alpha} + \mathfrak{B}^{\beta\alpha}$ of two free groups on a finite basis. We assume a $1:1$ correspondence between the generators of $\mathfrak{C}^{\beta}$ and $\mathfrak{C}^{\beta\alpha}$ and between the generators of $\mathfrak{B}^{\beta\alpha}$ and the permanent generators of $\mathfrak{B}^{\beta}$. The homomorphism $\mathfrak{K}^{\beta} \rightarrow \mathfrak{K}^{\beta\alpha}$ is given by this correspondence

$$\tag{10.5} \begin{array}{cc|c|l}
 & \mathfrak{C}^{\beta\alpha} & \mathfrak{B}^{\beta\alpha} & \\
\mathfrak{C}^{\beta} & 1 & 0 & \\
\hline
\mathfrak{B}^{\beta} & 0 & 1 & \text{perm.} \\
 & 0 & 0 & \text{temp.}
\end{array}$$

The homomorphism $\mathfrak{K}^{\beta\alpha} \rightarrow \mathfrak{K}^{\alpha}$ is given by

$$\tag{10.6} \begin{array}{cc|cc}
 & \mathfrak{C}^{\alpha} & \mathfrak{B}^{\alpha} & \\
\mathfrak{C}^{\beta\alpha} & W & 0 & Y \\
\hline
\mathfrak{B}^{\beta\alpha} & 0 & X & Z \\
 & & \text{perm.} & \text{temp.}
\end{array}.$$

where the submatrices W, X, Y and Z are those of (10.4). It is clear that the product of the homomorphisms $\mathfrak{K}^{\beta} \to \mathfrak{K}^{\beta\alpha} \to \mathfrak{K}^{\alpha}$ is the homomorphism $\mathfrak{K}^{\beta} \to \mathfrak{K}^{\alpha}$. We make of the enlarged collection of groups an inverse system as follows: If $\mathfrak{K}^{\beta\alpha} \to \mathfrak{K}^{\alpha} \to \mathfrak{K}^{\gamma}$, we define $\mathfrak{K}^{\beta\alpha} \to \mathfrak{K}^{\gamma}$ as the product. If $\mathfrak{K}^{\delta} \to \mathfrak{K}^{\beta} \to \mathfrak{K}^{\beta\alpha}$, we define $\mathfrak{K}^{\delta} \to \mathfrak{K}^{\beta\alpha}$ as the product. Then, if $\mathfrak{K}^{\epsilon\delta} \to \mathfrak{K}^{\beta} \to \mathfrak{K}^{\beta\alpha}$, we define $\mathfrak{K}^{\epsilon\delta} \to \mathfrak{K}^{\beta\alpha}$ as the product.

We shall define a boundary operation F mapping $\mathfrak{K}^{\beta\alpha}_{p+1}$ into $\mathfrak{K}_p^{\beta\alpha}$ and the latter into $\mathfrak{K}^{\beta\alpha}_{p-1}$. The boundary of any element of $\mathfrak{B}_q^{\beta\alpha}$ $(q = p, p+1)$ is defined to be zero. The boundary of a linear form in the basis of $\mathfrak{C}_q^{\beta\alpha}$ $(q = p, p+1)$ is the image in $\mathfrak{B}^{\beta\alpha}_{q-1}$ of the boundary of the same linear form in the corresponding basis elements of $\mathfrak{C}_q^{\beta}$. It is clear that the boundary operation commutes with the homomorphisms, and $\mathfrak{B}^{\beta\alpha}_{q-1}$ is the group of cycles which bound or have bounding multiples.

Let us consider the complete subsystem composed of those groups having two upper indices with all the homomorphisms between two such. In this system if $\mathfrak{K}^{\gamma\beta} \to \mathfrak{K}^{\alpha\delta}$ then $\mathfrak{C}^{\gamma\beta}$ images into $\mathfrak{C}^{\alpha\delta}$ and $\mathfrak{B}^{\gamma\beta}$ into $\mathfrak{B}^{\alpha\delta}$. This is proved by multiplying three matrices of the form (10.4), (10.5) and (10.6) respectively.

We introduce now the general topological group of coefficients $\mathfrak{G}$. $\mathfrak{K}^{\beta\alpha}(\mathfrak{G})$ is the group of linear forms in the basis elements of $\mathfrak{K}^{\beta\alpha}$ with coefficients from $\mathfrak{G}$. The boundary homomorphisms $\mathfrak{K}_p^{\beta\alpha}(\mathfrak{G}) \to \mathfrak{K}^{\beta\alpha}_{p-1}(\mathfrak{G})$ are defined in the usual way. Thus in $\mathfrak{K}_p^{\beta\alpha}(\mathfrak{G})$ we can distinguish a subgroup of cycles and a subgroup of bounding cycles, the latter being contained in the former. $\mathfrak{H}_p^{\beta\alpha}(\mathfrak{G})$ is obtained by reducing the group of cycles modulo the closure of the group of bounding cycles. A homomorphism $\mathfrak{K}_p^{\gamma\beta} \to \mathfrak{K}_p^{\alpha\delta}$ defines a homomorphism $\mathfrak{H}_p^{\gamma\beta}(\mathfrak{G}) \to \mathfrak{H}_p^{\alpha\delta}(\mathfrak{G})$ in the usual way (F commutes with $\mathfrak{K}^{\gamma\beta} \to \mathfrak{K}^{\alpha\delta}$). Thus $\{\mathfrak{H}_p^{\alpha\beta}(\mathfrak{G})\}$ is an inverse system. Furthermore it is equivalent with the system $\{\mathfrak{H}_p^{\alpha}(\mathfrak{G})\}$; for $\{\mathfrak{K}_q^{\alpha}\}$ and $\{\mathfrak{K}_q^{\alpha\beta}\}$ $(q = p-1, p, p+1)$ are equivalent and F commutes with the homomorphisms of the enlarged system containing both.

Now $\mathfrak{H}_p^{\alpha\beta}(\mathfrak{G})$ is the direct sum of a torsion group $\mathfrak{T}_p^{\alpha\beta}(\mathfrak{G})$ (elements representable by linear forms in basis elements of $\mathfrak{B}_p^{\alpha\beta}$) and a group $\mathfrak{S}_p^{\alpha\beta}(\mathfrak{G})$ of elements representable as linear forms in basis elements of $\mathfrak{C}_p^{\alpha\beta}$. Since $\mathfrak{K}^{\gamma\beta} \to \mathfrak{K}^{\alpha\delta}$ implies $\mathfrak{C}^{\gamma\beta} \to \mathfrak{C}^{\alpha\delta}$ and $\mathfrak{B}^{\gamma\beta} \to \mathfrak{B}^{\alpha\delta}$, $\{\mathfrak{T}_p^{\alpha\beta}(\mathfrak{G})\}$ and $\{\mathfrak{S}_p^{\alpha\beta}(\mathfrak{G})\}$ are inverse systems. Then their limit groups $\mathfrak{T}_p(\mathfrak{G})$ and $\mathfrak{S}_p(\mathfrak{G})$ respectively are subgroups of the limit group $\mathfrak{H}_p(\mathfrak{G})$ of $\{\mathfrak{H}_p^{\alpha\beta}(\mathfrak{G})\}$ and it is trivial that $\mathfrak{H}_p(\mathfrak{G})$ is their direct sum. This completes the proof.

REMARK. Let us note that the canonical system of chains $\{\mathfrak{K}_p^{\alpha\beta}\}$ can be defined for any finite number of successive values of p, say $0 \leqq p \leqq n$, so that

the groups of the two systems $\{\mathfrak{K}_p^{\alpha\beta}\}$ and $\{\mathfrak{K}_{p+1}^{\alpha\beta}\}$ are in 1:1 correspondence according to their upper indices and there is defined a homomorphism F mapping $\mathfrak{K}_{p+1}^{\alpha\beta}$ into $\mathfrak{K}_p^{\alpha\beta}$ which commutes with the homomorphisms and satisfies $FF=0$. If we pass to the corresponding systems over $\mathfrak{G}$, the group $\mathfrak{H}_p(\mathfrak{G})$ $(0\leqq p\leqq n-1)$ defined above is the p-th homology group over $\mathfrak{G}$ of $A \bmod B$.

III. A universal coefficient group.

11. The construction. We shall begin with an arbitrary bicompact group $\mathfrak{H}$, and we shall construct in terms of $\mathfrak{H}$ and the group $\mathfrak{G}$ two topological groups $\mathfrak{T}$ and $\mathfrak{S}$. Assuming that $\mathfrak{H}=\mathfrak{H}_p(A,B,\mathfrak{X})$ for a topological space A and closed subset B, we prove that $\mathfrak{T}=\mathfrak{T}_{p-1}(A,B,\mathfrak{G})$ and $\mathfrak{S}=\mathfrak{S}_p(A,B,\mathfrak{G})$ (see Theorem 10.1). The construction of $\mathfrak{T}$ and $\mathfrak{S}$ is not invariant since it involves a number of choices. To prove that $\mathfrak{T}$ and $\mathfrak{S}$ are invariants of the pair of groups $\mathfrak{H}$, $\mathfrak{G}$ it is necessary to carry along with the construction a certain amount of algebraic argument.

By Theorem 7.3, there is an inverse system of elementary groups having $\mathfrak{H}$ as limit group. Let $S=\{\mathfrak{H}^\alpha\}$ be any such system. As $\mathfrak{H}^\alpha$ is an elementary group, we may express it as a direct sum $\mathfrak{C}^\alpha+\mathfrak{D}^\alpha$ of a finite group $\mathfrak{D}^\alpha$ and a toral group $\mathfrak{C}^\alpha$ of finite dimension $d(\alpha)$.

Let us express $\mathfrak{D}^\alpha$ as the direct sum of finite cyclic groups $\mathfrak{X}'^{\alpha i}$ $(i=1,\cdots,\tau(\alpha))$ where $\theta^{\alpha i}$ is the order of $\mathfrak{X}'^{\alpha i}$. We shall suppose there is a fixed isomorphism between $\mathfrak{X}'^{\alpha i}$ and the subgroup of $\mathfrak{X}$ of order $\theta^{\alpha i}$. An element $\mathfrak{x}\in\mathfrak{X}$ $(\theta^{\alpha i}\mathfrak{x}=0)$ has a correspondent in $\mathfrak{X}'^{\alpha i}$ which is denoted $\mathfrak{x}\mathfrak{X}'^{\alpha i}$.

Let us express $\mathfrak{C}^\alpha$ as the direct sum of circular groups $\mathfrak{X}^{\alpha i}$ $(i=1,\cdots,d(\alpha))$. We choose a fixed continuous isomorphism between $\mathfrak{X}$ and $\mathfrak{X}^{\alpha i}$ and denote the correspondent of $\mathfrak{x}\in\mathfrak{X}$ by $\mathfrak{x}\mathfrak{X}^{\alpha i}$.

Let $\tilde{\mathfrak{X}}$ be a continuous homomorphic image of $\mathfrak{X}$ in $\mathfrak{C}^\alpha$. The image of $\mathfrak{x}\in\mathfrak{X}$ is denoted $\mathfrak{x}\tilde{\mathfrak{X}}$. For some elements $\mathfrak{x}_i(\mathfrak{x})$, we must have [9] $\mathfrak{x}\tilde{\mathfrak{X}}=\mathfrak{x}_i(\mathfrak{x})\mathfrak{X}^{\alpha i}$. We prove that there are unique integers u_i such that

$$\mathfrak{x}\tilde{\mathfrak{X}}=\mathfrak{x}u_i\mathfrak{X}^{\alpha i}.$$

The function $\mathfrak{x}_i(\mathfrak{x})$ is a continuous homomorphism of $\mathfrak{X}$ into itself (this follows from the definition of direct sum). It is well known that any such character of $\mathfrak{X}$ is given by an integer u_i satisfying $\mathfrak{x}_i(\mathfrak{x})=u_i\mathfrak{x}$ for each $\mathfrak{x}\in\mathfrak{X}$.

The subgroup $\mathfrak{C}^\alpha$ of $\mathfrak{H}^\alpha$ is uniquely determined since it is the component of zero. The group $\mathfrak{D}^\alpha$ may in general be chosen in several ways. Suppose $\tilde{\mathfrak{D}}^\alpha$ is a finite subgroup so that $\mathfrak{H}^\alpha=\mathfrak{C}^\alpha+\tilde{\mathfrak{D}}^\alpha$. Let us decompose $\tilde{\mathfrak{D}}^\alpha$ into cyclic subgroups $\tilde{\mathfrak{X}}'^{\alpha i}$; and let us decompose $\mathfrak{C}^\alpha$ in some new way into circular subgroups $\tilde{\mathfrak{X}}^{\alpha i}$. Then the element $\mathfrak{x}\tilde{\mathfrak{X}}'^{\alpha i}$ $(\mathfrak{x}\in\mathfrak{X}, \tilde{\theta}^{\alpha i}\mathfrak{x}=0)$ may be written

$$\mathfrak{x}\bar{\mathfrak{X}}'^{\alpha i} = \mathfrak{x}_j{}^i(\mathfrak{x})\mathfrak{X}^{\alpha j} + \mathfrak{x}'_j{}^i(\mathfrak{x})\mathfrak{X}'^{\alpha j}.$$

We shall prove that there are integers $v_j{}^i$ and $w_j{}^i$ which are unique mod $\bar{\theta}^{\alpha i}$ such that

$$\mathfrak{x}\bar{\mathfrak{X}}'^{\alpha i} = \mathfrak{x}v_j{}^i\mathfrak{X}^{\alpha j} + \mathfrak{x}w_j{}^i\mathfrak{X}'^{\alpha j}.$$

The function $\mathfrak{x}_j{}^i(\mathfrak{x})$ maps homomorphically the subgroup of $\mathfrak{X}$ with generator $1/\bar{\theta}^{\alpha i}$ into $\mathfrak{X}$ and therefore into a subgroup of itself. Thus there is an integer $v_j{}^i$ unique mod $\bar{\theta}^{\alpha i}$ such that $1/\bar{\theta}^{\alpha i}$ is imaged into $v_j{}^i$ times itself. In the same way we obtain $w_j{}^i$.

Thus between the two decompositions of $\mathfrak{H}^\alpha$ into the direct sums of circular groups and finite cyclic groups, we have the transformations

$$\left\{\begin{array}{l} \mathfrak{x}\bar{\mathfrak{X}}^{\alpha i} = \mathfrak{x}u_i{}^{\alpha i}\mathfrak{X}^{\alpha j} \\ \mathfrak{x}\bar{\mathfrak{X}}'^{\alpha i} = \mathfrak{x}v_j{}^{\alpha i}\mathfrak{X}^{\alpha j} + \mathfrak{x}w_j{}^{\alpha i}\mathfrak{X}'^{\alpha j}, \end{array}\right. \tag{11.1}$$

$$\left\{\begin{array}{l} \mathfrak{x}\mathfrak{X}^{\alpha k} = \mathfrak{x}\bar{u}_i{}^{\alpha k}\bar{\mathfrak{X}}^{\alpha i} \\ \mathfrak{x}\mathfrak{X}'^{\alpha k} = \mathfrak{x}\bar{v}_i{}^{\alpha k}\bar{\mathfrak{X}}^{\alpha i} + \mathfrak{x}\bar{w}_i{}^{\alpha k}\bar{\mathfrak{X}}'^{\alpha i}. \end{array}\right. \tag{11.2}$$

As each transformation is the inverse of the other we obtain

$$\left\{\begin{array}{l} \bar{u}_j{}^{\alpha i}u_k{}^{\alpha j} = \delta_k{}^i \quad (\delta_i{}^i = 1,\ \delta_k{}^i = 0\ \ (i \neq k)), \\ \bar{v}_j{}^{\alpha i}u_k{}^{\alpha j} + \bar{w}_j{}^{\alpha i}v_k{}^{\alpha j} \equiv 0 \qquad (\text{mod } \theta^{\alpha i}), \\ \bar{w}_j{}^{\alpha i}w_k{}^{\alpha j} \equiv \delta_k{}^i \quad (\text{mod } \theta^{\alpha i}). \end{array}\right. \tag{11.3}$$

If $\mathfrak{H}^\beta$ is a refinement of $\mathfrak{H}^\alpha$, $H_\beta{}^\alpha$ maps the circular subgroups $\mathfrak{X}^{\beta i}$ of $\mathfrak{H}^\beta$ into such in $\mathfrak{H}^\alpha$, and maps the finite groups $\mathfrak{X}'^{\beta i}$ into such in $\mathfrak{H}^\alpha$. Therefore we may adopt the notation just described and write $H_\beta{}^\alpha$ in the form

$$\left\{\begin{array}{l} H_\beta{}^\alpha(\mathfrak{x}\mathfrak{X}^{\beta i}) = \mathfrak{x}x_{\beta j}^{\alpha i}\mathfrak{X}^{\alpha j} \\ H_\beta{}^\alpha(\mathfrak{x}\mathfrak{X}'^{\beta i}) = \mathfrak{x}y_{\beta j}^{\alpha i}\mathfrak{X}^{\alpha j} + \mathfrak{x}z_{\beta j}^{\alpha i}\mathfrak{X}'^{\alpha j} \end{array}\right. \tag{11.4}$$

where the integers $x_{\beta j}^{\alpha i}$ are uniquely determined, and the integers $y_{\beta j}^{\alpha i}$ and $z_{\beta j}^{\alpha i}$ are unique mod $\theta^{\beta i}$.

Similarly for the other bases

$$\left\{\begin{array}{l} H_\beta{}^\alpha(\mathfrak{x}\bar{\mathfrak{X}}^{\beta i}) = \mathfrak{x}\bar{x}_{\beta j}^{\alpha i}\bar{\mathfrak{X}}^{\alpha j} \\ H_\beta{}^\alpha(\mathfrak{x}\bar{\mathfrak{X}}'^{\beta i}) = \mathfrak{x}\bar{y}_{\beta j}^{\alpha i}\bar{\mathfrak{X}}^{\alpha j} + \mathfrak{x}\bar{z}_{\beta j}^{\alpha i}\bar{\mathfrak{X}}'^{\alpha j}. \end{array}\right. \tag{11.5}$$

If we apply successively the transformations (11.2) for $\mathfrak{H}^\beta$ then (11.5) and finally (11.1), we must obtain the transformation (11.4). We have therefore the relations

$$\left\{\begin{array}{l} \bar{u}_j{}^{\beta i}\bar{x}_{\beta k}^{\alpha j}u_l{}^{\alpha k} = x_{\beta l}^{\alpha i} \\ \bar{v}_j{}^{\beta i}\bar{x}_{\beta k}^{\alpha j}u_l{}^{\alpha k} + \bar{w}_j{}^{\beta i}\bar{y}_{\beta k}^{\alpha j}u_l{}^{\alpha k} + \bar{w}_j{}^{\beta i}\bar{z}_{\beta k}^{\alpha j}v_l{}^{\alpha k} \equiv y_{\beta l}^{\alpha i} \quad (\text{mod } \theta^{\beta i}) \\ \bar{w}_j{}^{\beta i}\bar{z}_{\beta k}^{\alpha j}w_l{}^{\alpha k} \equiv z_{\beta l}^{\alpha i} \quad (\text{mod } \theta^{\beta i}) \end{array}\right. \tag{11.6}$$

where of course there is no summation on β.

In terms of $\mathfrak{G}$, $\{\mathfrak{H}^\alpha\}$ and the decomposition of $\mathfrak{H}^\alpha$ into subgroups $\mathfrak{X}^{\alpha i}$ and $\mathfrak{X}'^{\alpha i}$, we construct an inverse system $\{\mathfrak{T}^\alpha\}$ as follows. Let $\mathfrak{G}^{*\alpha i}$ $(i=1,\cdots,\tau(\alpha))$ be a group bicontinuously isomorphic with the group obtained by reducing $\mathfrak{G}$ modulo the closure of the subgroup of elements divisible by $\theta^{\alpha i}$. Under this fixed homomorphism of $\mathfrak{G}$ into $\mathfrak{G}^{*\alpha i}$, denote the image of $\mathfrak{g}\in\mathfrak{G}$ by $\mathfrak{g}\mathfrak{G}^{*\alpha i}$. Let $\mathfrak{T}^\alpha$ be the direct sum $\sum_{i=1}^{\tau(\alpha)}\mathfrak{G}^{*\alpha i}$. The number $z^{\alpha i}_{\beta j}$ of (11.4) is such that $\mathfrak{x} z^{\alpha i}_{\beta j}\theta^{\alpha j}$ (not summed on j) is the zero of $\mathfrak{X}$ if $\theta^{\beta i}\mathfrak{x}=0$. If we let $\mathfrak{x}$ be the element $1/\theta^{\beta i}$, we see that there is an integer $s^{\alpha i}_{\beta j}$ such that

$$s^{\alpha i}_{\beta j}\theta^{\beta i}=z^{\alpha i}_{\beta j}\theta^{\alpha j}\quad(\text{not summed on } j \text{ and } \beta). \tag{11.7}$$

We define a homomorphism $H_\beta{}^\alpha$ of $\mathfrak{T}^\beta$ into $\mathfrak{T}^\alpha$ by the relations

$$H_\beta{}^\alpha(\mathfrak{g}\mathfrak{G}^{*\beta i})=\mathfrak{g}s^{\alpha i}_{\beta j}\mathfrak{G}^{*\alpha j}\qquad(i=1,\cdots,\tau(\beta)). \tag{11.8}$$

Since $z^{\alpha i}_{\beta j}$ is unique mod $\theta^{\beta i}$, $s^{\alpha i}_{\beta j}$ is unique mod $\theta^{\alpha j}$, thus the right side is a unique element of $\mathfrak{T}^\alpha$.

Due to the ambiguity in representing an element of $\mathfrak{G}^{*\beta i}$, we must prove that $H_\beta{}^\alpha$ is uniquely defined. Suppose $\mathfrak{g}\mathfrak{G}^{*\beta i}=\mathfrak{g}'\mathfrak{G}^{*\beta i}$, then $\mathfrak{g}-\mathfrak{g}'$ is a limit of elements of $\mathfrak{G}$ divisible by $\theta^{\beta i}$. Therefore $(\mathfrak{g}-\mathfrak{g}')s^{\alpha i}_{\beta j}$ is a limit of elements divisible by $s^{\alpha i}_{\beta j}\theta^{\beta i}$. By (11.7), $(\mathfrak{g}-\mathfrak{g}')s^{\alpha i}_{\beta j}$ is a limit of elements divisible by $\theta^{\alpha j}$. Therefore $\mathfrak{g}s^{\alpha i}_{\beta j}$ and $\mathfrak{g}'s^{\alpha i}_{\beta j}$ image into the same element of $\mathfrak{G}^{*\alpha j}$, and $H_\beta{}^\alpha$ is unique.

To see that $H_\beta{}^\alpha$ is continuous, let U be a neighborhood of zero in $\mathfrak{T}^\alpha$. Then U is given as the product space of neighborhoods U^j $(j=1,\cdots,\tau(\alpha))$ of neighborhoods of zero in the $\mathfrak{G}^{*\alpha j}$. Let V be a neighborhood of zero in $\mathfrak{G}$ such that V is imaged into U^j $(j=1,\cdots,\tau(\alpha))$ under the transformation sending $\mathfrak{g}\in\mathfrak{G}$ into $\mathfrak{g}s^{\alpha i}_{\beta j}\mathfrak{G}^{*\alpha j}$ (not summed on j). The image of V in $\mathfrak{G}^{*\beta i}$ under the homomorphism $\mathfrak{g}\to\mathfrak{g}\mathfrak{G}^{*\beta i}$ is a neighborhood U'^i of zero in $\mathfrak{G}^{*\beta i}$. The product space U' of the U'^i is a neighborhood of zero in $\mathfrak{T}^\beta$ mapped by $H_\beta{}^\alpha$ into U.

If the homomorphism $H_\beta{}^\alpha$ of $\mathfrak{T}^\beta$ into $\mathfrak{T}^\alpha$ be defined for every pair α, β of ordinals such that $\mathfrak{H}^\beta$ is a refinement of $\mathfrak{H}^\alpha$, we obtain an inverse system $\{\mathfrak{T}^\alpha\}$. (The verification of axiom c) of No. 2 is immediate. Since $\{\mathfrak{H}^\alpha\}$ satisfies axioms a) and b), we deduce corresponding properties for the integers x, y, and z of (11.4) from which it follows that a) and b) hold in $\{\mathfrak{T}^\alpha\}$). Let us prove that $\{\mathfrak{T}^\alpha\}$ is independent of the decomposition of $\mathfrak{H}^\alpha$ into subgroups. If $\mathfrak{H}^\alpha$ be decomposed into the subgroups $\bar{\mathfrak{X}}^{\alpha i}$ and $\bar{\mathfrak{X}}'^{\alpha i}$, we construct groups $\bar{\mathfrak{G}}^{*\alpha i}$ in a similar way and obtain a new group $\bar{\mathfrak{T}}^\alpha$. From (11.1) and (11.2) we deduce integers $t_j{}^{\alpha i}$ and $\bar{t}_j{}^{\alpha i}$ such that

$$\left.\begin{aligned} t_j{}^{\alpha i}\bar{\theta}^{\alpha i} &= w_j{}^{\alpha i}\theta^{\alpha j} \\ \bar{t}_j{}^{\alpha i}\theta^{\alpha i} &= \bar{w}_j{}^{\alpha i}\theta^{\alpha j} \end{aligned}\right\} \text{ (not summed on } j\text{).}$$

From these relations and (11. 3), we deduce

$$\bar{t}_j{}^{\alpha i}t_k{}^{\alpha j} \equiv \delta_k{}^i \pmod{\theta^{\alpha k}}.$$

Then the transformation

$$f^\alpha(\mathfrak{g}\bar{\mathfrak{G}}^{*\alpha i}) = \mathfrak{g}t_j{}^{\alpha i}\mathfrak{G}^{*\alpha j}$$

has the inverse

$$\bar{f}^\alpha(\mathfrak{g}\mathfrak{G}^{*\alpha j}) = \mathfrak{g}\bar{t}_k{}^{\alpha j}\bar{\mathfrak{G}}^{*\alpha k},$$

and is therefore a bicontinuous isomorphism between $\mathfrak{T}^\alpha$ and $\bar{\mathfrak{T}}^\alpha$. By (11. 5), there are integers $\bar{s}^{\alpha i}_{\beta j}$ satisfying

$$\bar{s}^{\alpha i}_{\beta j}\bar{\theta}^{\beta i} = \bar{z}^{\alpha i}_{\beta j}\bar{\theta}^{\alpha j} \quad \text{(not summed on } j\text{).}$$

By (11. 6) and (11. 7), we obtain

$$\bar{t}_i{}^{\beta h}\bar{s}^{\alpha i}_{\beta j}t_k{}^{\alpha j} \equiv s^{\alpha h}_{\beta k} \pmod{\theta^{\alpha k}}.$$

Therefore the homomorphism

$$\bar{H}_\beta{}^\alpha(\mathfrak{g}\bar{\mathfrak{G}}^{*\beta i}) = \mathfrak{g}s^{\alpha i}_{\beta j}\bar{\mathfrak{G}}^{*\alpha i}$$

carries over into the homomorphism $H_\beta{}^\alpha$ under the isomorphisms f^α: $H_\beta{}^\alpha = f^\alpha\bar{H}_\beta{}^\alpha\bar{f}^\beta$. Thus the inverse system $\{\bar{\mathfrak{T}}^\alpha\}$ is isomorphic with $\{\mathfrak{T}^\alpha\}$. It follows that $\{\mathfrak{T}^\alpha\}$ is uniquely defined by $\mathfrak{G}$ and $\{\mathfrak{H}^\alpha\}$.

A complete subsystem of $S = \{\mathfrak{H}^\alpha\}$ defines in the same way a complete subsystem of $\{\mathfrak{T}^\alpha\}$. Thus equivalent systems S and $\bar{S}$ determine equivalent systems $\{\mathfrak{T}^\alpha\}$ and $\{\bar{\mathfrak{T}}^\alpha\}$. By Theorem 7. 4, it follows that $\{\mathfrak{T}^\alpha\}$ is determined up to equivalent systems by $\mathfrak{G}$ and $\mathfrak{H}$. Hence by Theorem 4. 1, the limit group $\mathfrak{T}$ of $\{\mathfrak{T}^\alpha\}$ is determined up to bicontinuous isomorphisms by $\mathfrak{G}$ and $\mathfrak{H}$.

In terms of $\mathfrak{G}$, $\{\mathfrak{H}^\alpha\}$, and the decomposition of $\mathfrak{H}^\alpha$ into subgroups $\mathfrak{X}^{\alpha i}$ and $\mathfrak{X}'^{\alpha i}$, we construct an inverse system $\{\mathfrak{S}^\alpha\}$ as follows. Let $\mathfrak{G}^{\alpha i}$ $(i = 1, \cdots, d(\alpha))$ be a group bicontinuously isomorphic with $\mathfrak{G}$. Let $\mathfrak{G}'^{\alpha i}$ $(i = 1, \cdots, \tau(\alpha))$ be a group bicontinuously isomorphic with the subgroup of elements of $\mathfrak{G}$ of orders dividing $\theta^{\alpha i}$. As usual the element of $\mathfrak{G}^{\alpha i}$ $(\mathfrak{G}'^{\alpha i})$ corresponding to $\mathfrak{g} \in \mathfrak{G}$ is denoted $\mathfrak{g}\mathfrak{G}^{\alpha i}$ $(\mathfrak{g}\mathfrak{G}'^{\alpha i})$. Let $\mathfrak{S}^\alpha$ be the direct sum

$$\mathfrak{S}^\alpha = \sum_{i=1}^{d(\alpha)} \mathfrak{G}^{\alpha i} + \sum_{i=1}^{\tau(\alpha)} \mathfrak{G}'^{\alpha i}.$$

Let the homomorphism $\check{H}_\beta{}^\alpha$ of $\mathfrak{S}^\beta$ into $\mathfrak{S}^\alpha$ be defined by

$$(11.9)\qquad \begin{cases} H_\beta{}^\alpha(\mathfrak{g}\mathfrak{G}^{\beta i}) = \mathfrak{g}x^{\alpha i}_{\beta j}\mathfrak{G}^{\alpha j} \\ H_\beta{}^\alpha(\mathfrak{g}\mathfrak{G}'^{\beta i}) = \mathfrak{g}y^{\alpha i}_{\beta j}\mathfrak{G}^{\alpha j} + \mathfrak{g}z^{\alpha i}_{\beta j}\mathfrak{G}'^{\alpha j} \end{cases}$$

where the integers x, y and z are those of (11.4). Since $y^{\alpha i}_{\beta j}$ and $z^{\alpha i}_{\beta j}$ are unique mod $\theta^{\beta i}$ and $\mathfrak{g}$ in the second equation is of order $\theta^{\beta i}$, the transformation is uniquely defined. It is verified without difficulty that $\{\mathfrak{S}^\alpha\}$ is an inverse system.

Let us prove that $\{\mathfrak{S}^\alpha\}$ is independent of the decomposition of $\mathfrak{H}^\alpha$ into subgroups. If $\mathfrak{H}^\alpha$ be decomposed into the subgroups $\bar{\mathfrak{X}}^{\alpha i}$ and $\bar{\mathfrak{X}}'^{\alpha i}$, we construct groups $\bar{\mathfrak{G}}^{\alpha i}$ and $\bar{\mathfrak{G}}'^{\alpha i}$ in a similar way, and obtain a new group $\bar{\mathfrak{S}}^\alpha$. Using (11.1), the transformation f^α defined by

$$f^\alpha(\mathfrak{g}\bar{\mathfrak{G}}^{\alpha i}) = \mathfrak{g}u_j{}^{\alpha i}\mathfrak{G}^{\alpha j}$$
$$f^\alpha(\mathfrak{g}\bar{\mathfrak{G}}'^{\alpha i}) = \mathfrak{g}v_j{}^{\alpha i}\mathfrak{G}^{\alpha j} + \mathfrak{g}w_j{}^{\alpha i}\mathfrak{G}'^{\alpha\alpha j}$$

has the inverse

$$\bar{f}^\alpha(\mathfrak{g}\mathfrak{G}^{\alpha j}) = \mathfrak{g}\bar{u}_k{}^{\alpha j}\bar{\mathfrak{G}}^{\alpha k}$$
$$\bar{f}^\alpha(\mathfrak{g}\mathfrak{G}'^{\alpha j}) = \mathfrak{g}\bar{v}_k{}^{\alpha j}\bar{\mathfrak{G}}^{\alpha k} + \mathfrak{g}\bar{w}_k{}^{\alpha j}\bar{\mathfrak{G}}'^{\alpha k}.$$

This follows from the relations (11.3). And the homomorphism $\bar{H}_\beta{}^\alpha$ defined by

$$\bar{H}_\beta{}^\alpha(\mathfrak{g}\bar{\mathfrak{G}}^{\beta i}) = \mathfrak{g}\bar{x}^{\alpha i}_{\beta j}\bar{\mathfrak{G}}^{\alpha j}$$
$$\bar{H}_\beta{}^\alpha(\mathfrak{g}\bar{\mathfrak{G}}'^{\beta i}) = \mathfrak{g}\bar{y}^{\alpha i}_{\beta j}\bar{\mathfrak{G}}^{\alpha j} + \mathfrak{g}\bar{z}^{\alpha i}_{\beta j}\bar{\mathfrak{G}}'^{\alpha j}$$

satisfies, in view of (11.6), the relation $H_\beta{}^\alpha = f^\alpha\bar{H}_\beta{}^\alpha\bar{f}^\beta$. And we see that $\{\mathfrak{S}^\alpha\}$ is determined up to isomorphic systems by $\{\mathfrak{H}^\alpha\}$ and $\mathfrak{G}$.

Just as above we can prove that the limit group $\mathfrak{S}$ of $\{\mathfrak{S}^\alpha\}$ is determined up to bicontinuous isomorphisms by $\mathfrak{G}$ and $\mathfrak{H}$.

12. The isomorphism. We now assume that $\mathfrak{H} = \mathfrak{H}_p(A, B, \mathfrak{X})$ for a topological space A and closed subset B. We shall prove that $\mathfrak{T} = \mathfrak{T}_{p-1}(A, B, \mathfrak{G})$ and $\mathfrak{S} = \mathfrak{S}_p(A, B, \mathfrak{G})$.

We suppose that $\{\mathfrak{K}_q{}^\alpha\}$ $(q = 0, 1, \cdots, p+1)$ is the canonical inverse system of chains constructed in No. 10 for the space A. We suppose, for each α, bases have been chosen in the groups $\mathfrak{K}_p{}^\alpha$ and $\mathfrak{K}^\alpha{}_{p-1}$ so that the boundary relations for the chains of $\mathfrak{K}_p{}^\alpha$ are in quasi-canonical form. We suppose in particular that we have split these basis elements into five classes so that the relations (8.4) hold.

If we pass to the corresponding homology groups over $\mathfrak{G}$, we find by Theorem 8 that

$$\mathfrak{T}^\alpha{}_{p-1}(\mathfrak{G}) = \sum \mathfrak{G}^{*\alpha i} \qquad (i = 1, \cdots, \tau_{p-1}(\alpha))$$

where $\mathfrak{G}^{*\alpha i}$ is obtained by reducing $\mathfrak{G}$ modulo the closure of the subgroup of elements divisible by $\theta^{\alpha i}_{p-1}$, and

$$\mathfrak{S}_p{}^\alpha(\mathfrak{G}) = \sum \mathfrak{G}^{\alpha j} + \sum \mathfrak{G}'^{\alpha k} \quad (j = 1, \cdots, R_p(\alpha)\,;\ k = 1, \cdots, \tau_{p-1}(\alpha))$$

where $\mathfrak{G}^{\alpha j}$ is isomorphic with $\mathfrak{G}$ and $\mathfrak{G}'^{\alpha k}$ with the subgroup of $\mathfrak{G}$ of order $\theta^{\alpha k}_{p-1}$. In particular

$$\mathfrak{H}_p{}^\alpha(\mathfrak{X}) = \sum \mathfrak{X}^{\alpha j} + \sum \mathfrak{X}'^{\alpha k}. \tag{12.1}$$

Since $\mathfrak{H}_p(A, B, \mathfrak{X})$ is the limit group of $\{\mathfrak{H}_p{}^\alpha(\mathfrak{X})\}$, we may suppose that the latter is the system used in No. 11 for constructing $\mathfrak{T}$ and $\mathfrak{S}$, and we may suppose moreover that the decomposition there used of each $\mathfrak{H}_p{}^\alpha(\mathfrak{X})$ into a direct sum is the decomposition (12.1). It is then immediately clear that $\mathfrak{T}^\alpha = \mathfrak{T}^\alpha{}_{p-1}(\mathfrak{G})$ and $\mathfrak{S}^\alpha = \mathfrak{S}_p{}^\alpha(\mathfrak{G})$. These isomorphisms are set up in an obvious way. We have only to prove that under these isomorphisms the homomorphisms of the systems $\{\mathfrak{T}^\alpha\}$ and $\{\mathfrak{S}^\alpha\}$ carry over into those of $\{\mathfrak{T}^\alpha{}_{p-1}(\mathfrak{G})\}$ and $\{\mathfrak{S}_p{}^\alpha(\mathfrak{G})\}$ respectively. The latter case is trivial. For the former, suppose the homomorphism of $\mathfrak{K}_p{}^\beta$ into $\mathfrak{K}_p{}^\alpha$ has on the basis element $d_p{}^{\beta k}$ of class four (see (8.4)) the value

$$y^{\alpha k}_{\beta j} c_p{}^{\alpha j} + z^{\alpha k}_{\beta j} d_p{}^{\alpha j} + w^{\alpha k}_{\beta j} e_p{}^{\alpha j},$$

(we recall that, in a canonical system of chains, $\mathfrak{C}_p{}^\beta$ is mapped into $\mathfrak{C}_p{}^\alpha$). Then, for the homomorphism of $\mathfrak{H}_p{}^\beta(\mathfrak{X})$ into $\mathfrak{H}_p{}^\alpha(\mathfrak{X})$,

$$H_\beta{}^\alpha(\mathfrak{x}\mathfrak{X}'^{\beta k}) = \mathfrak{x} y^{\alpha k}_{\beta j} \mathfrak{X}^{\alpha j} + \mathfrak{x} z^{\alpha k}_{\beta j} \mathfrak{X}'^{\alpha j}$$

As $FH = HF$, we have

$$\theta^{\beta k}_{p-1} H_\beta{}^\alpha(b^{\beta k}_{p-1}) = H_\beta{}^\alpha F(d_p{}^{\beta k}) = FH_\beta{}^\alpha(d_p{}^{\beta k}) = z^{\alpha k}_{\beta j}\theta^{\alpha j}_{p-1} b^{\alpha j}_{p-1} + w^{\alpha k}_{\beta j} a^{\alpha j}_{p-1}.$$

It follows that $z^{\alpha k}_{\beta j}\theta^{\alpha j}_{p-1}$ (not summed on j) is divisible by $\theta^{\beta k}_{p-1}$. If $s^{\alpha k}_{\beta j}$ denotes the quotient, then

$$H_\beta{}^\alpha(b^{\beta k}_{p-1}) \sim s^{\alpha k}_{\beta j} b^{\alpha j}_{p-1}.$$

Hence

$$H_\beta{}^\alpha(\mathfrak{g}\mathfrak{G}^{*\beta k}) = \mathfrak{g} s^{\alpha k}_{\beta j} \mathfrak{G}^{*\alpha j}.$$

As this is the homomorphism (11.8), we have proved

THEOREM 12. *The group $\mathfrak{H}_p(A, B, \mathfrak{G})$ is the direct sum $\mathfrak{T}_p + \mathfrak{S}_p$ where $\mathfrak{T}_p$ is an invariant of the pair $\mathfrak{G}$, $\mathfrak{H}_{p+1}(A, B, \mathfrak{X})$ and $\mathfrak{S}_p$ is an invariant of the pair $\mathfrak{G}$, $\mathfrak{H}_p(A, B, \mathfrak{X})$. Thus the group $\mathfrak{X}$ is a universal coefficient group for the Čech homology theory of a topological space.*

It is well known that in a compact metric space the Čech theory and the Vietoris [18] theory are equivalent; hence

COROLLARY. *The group $\mathfrak{X}$ is universal for the Vietoris homology theory of a compact metric space.*

REMARK. Čech has pointed out to me the following simple invariant definition of the group $\mathfrak{S}$ in terms of the groups $\mathfrak{G}$ and $\mathfrak{H}$. Let $\mathfrak{H}^*$ be the discrete group of continuous characters of $\mathfrak{H}$. Then $\mathfrak{S}$ is defined to be the group of all homomorphic mappings of $\mathfrak{H}^*$ into $\mathfrak{G}$. If U is a neighborhood of zero in $\mathfrak{G}$ and $h_1, h_2, \cdots, h_k$ are a finite number of elements of $\mathfrak{H}^*$, the set V of elements of $\mathfrak{S}$ mapping each h_i into U is called a neighborhood of zero in $\mathfrak{S}$. That $\mathfrak{S}$ is the group constructed in No. 11 is obvious if $\mathfrak{H}$ is an elementary group. Otherwise let $\mathfrak{H}$ be a limit group of the inverse system $\{\mathfrak{H}^a\}$ of elementary groups. Let $\mathfrak{S}^a$ be the group of homomorphic mappings of $\mathfrak{H}^{*a}$ into $\mathfrak{G}$. As $\{\mathfrak{H}^{*a}\}$ is a direct system (No. 5), it is easy to see that $\{\mathfrak{S}^a\}$ is an inverse system and that its limit group is the group of homomorphic mappings of $\mathfrak{H}^*$ into $\mathfrak{G}$. In this way the case of a general bicompact group $\mathfrak{H}$ is reduced to that of an elementary group. If one could find an equally simple invariant definition of $\mathfrak{T}$ in terms of $\mathfrak{G}$ and $\mathfrak{H}$, the argument of the preceding two sections could be greatly simplified.

IV. The infinite complex.

13. Infinite cycles. Let K be an infinite complex with a countable number of cells. We require that K be locally finite in the sense that the star of any vertex is finite. Let L be a closed subcomplex of K and let the p-cells of $K - L$ be ordered in a sequence: $E_p{}^i$ $(i = 1, 2, \cdots)$. A p-chain over $\mathfrak{G}$ of K mod L is an infinite linear form: $\mathfrak{g}_i E_p{}^i$. These constitute a group $\mathfrak{K}_p(K, L, \mathfrak{G})$ in which a topology is introduced as follows. If U is a neighborhood of zero in $\mathfrak{G}$ and n is an integer, the set V of those chains whose first n coefficients lie in U is a neighborhood of zero in $\mathfrak{K}$. As in the case of a finite complex, the boundary operator F is a continuous homomorphism of $\mathfrak{K}_p$ into $\mathfrak{K}_{p-1}$. Then cycles and bounding cycles can be distinguished. The group $\mathfrak{H}_p(K, L, \mathfrak{G})$ is obtained by reducing the group of cycles modulo the closure of the subgroup of bounding cycles.

14. Universal group theorem. If we can show that the homology theory of an infinite complex can be analyzed in terms of inverse systems of homology groups of finite complexes, the proof that $\mathfrak{X}$ is a universal coefficient group for infinite cycles is clearly contained in the preceding sections. The definition just given of $\mathfrak{H}_p(K, L, \mathfrak{G})$ is standard since it follows closely the spirit of the definition of Lefschetz [15; p. 299]. Consider the following alternative definition.

Let $\{L^\alpha\}$ be a sequence of closed subcomplexes of K such that $\prod L^\alpha = L$, $L^\alpha \supset L^{\alpha+1}$, and each $K - L^\alpha$ is finite. Let $\mathfrak{K}_p{}^\alpha(\mathfrak{G}) = \mathfrak{K}_p(K, L^\alpha, \mathfrak{G})$. Suppose $L^\alpha \supset L^\beta$. To each chain of $\mathfrak{K}_p{}^\beta(\mathfrak{G})$ we associate the chain of $\mathfrak{K}_p{}^\alpha(\mathfrak{G})$ obtained by omitting its terms involving cells on L^α. This is a homomorphism of $\mathfrak{K}_p{}^\beta(\mathfrak{G})$ into $\mathfrak{K}_p{}^\alpha(\mathfrak{G})$. It preserves cycles and bounding cycles, and therefore induces a continuous homomorphism of $\mathfrak{H}_p{}^\beta(\mathfrak{G})$ into $\mathfrak{H}_p{}^\alpha(\mathfrak{G})$. In this way $\{\mathfrak{H}_p{}^\alpha(\mathfrak{G})\}$ constitutes an inverse system with a limit group $\mathfrak{H}'_p(K, L, \mathfrak{G})$. It is clear that $\mathfrak{H}'_p(\mathfrak{G})$ does not depend on the particular sequence $\{L^\alpha\}$; for any other sequence $\{L'^\alpha\}$ satisfying the same conditions determines an equivalent system $\{\mathfrak{H}'_p{}^\alpha(\mathfrak{G})\}$. We will see under what conditions the group $\mathfrak{H}'_p(K, L, \mathfrak{G})$ is $\mathfrak{H}_p(K, L, \mathfrak{G})$.

We define a continuous homomorphism of $\mathfrak{H}_p(\mathfrak{G})$ into $\mathfrak{H}'_p(\mathfrak{G})$ as follows. If $\mathfrak{c}_p$ is an infinite cycle mod L, let $\mathfrak{c}_p{}^\alpha$ be the cycle mod L^α obtained by omitting terms of $\mathfrak{c}_p$ involving cells on L^α. Suppose $\mathfrak{h} \in \mathfrak{H}_p(\mathfrak{G})$ has $\mathfrak{c}_p$ as a representative cycle. Then $\mathfrak{c}_p{}^\alpha$ is representative of a class $\mathfrak{h}^\alpha \in \mathfrak{H}_p{}^\alpha(\mathfrak{G})$. Clearly $\{\mathfrak{h}^\alpha\}$ is an element $\mathfrak{h}' \in \mathfrak{H}'_p(\mathfrak{G})$. If $\mathfrak{c}'_p \sim \mathfrak{c}_p$, then $\mathfrak{c}'_p{}^\alpha \sim \mathfrak{c}_p{}^\alpha$, and $\mathfrak{h}'$ is independent of the representative $\mathfrak{c}_p$ of $\mathfrak{h}$. Thus $f(\mathfrak{h}) = \mathfrak{h}'$ is a continuous homomorphism of $\mathfrak{H}_p(\mathfrak{G})$ into $\mathfrak{H}'_p(\mathfrak{G})$. In fact f is an isomorphism of $\mathfrak{H}_p(\mathfrak{G})$ into a subgroup of $\mathfrak{H}'_p(\mathfrak{G})$. For suppose $f(\mathfrak{h}) = 0$. Then each $\mathfrak{c}_p{}^\alpha$ is a limit of bounding cycles mod L^α. That is: to a neighborhood V of $\mathfrak{c}_p{}^\alpha$ corresponds a chain $\mathfrak{k}_{p+1}$ such that $F(\mathfrak{k}_{p+1})$ mod L^α lies in V. If W is a neighborhood of $\mathfrak{c}_p$ determined by U in $\mathfrak{G}$ and the integer n, choose L^α so that it does not contain the first n p-cells of $K - L$. Let V be the neighborhood of $\mathfrak{c}_p{}^\alpha$ determined by U in $\mathfrak{G}$. Then $F(\mathfrak{k}_{p+1})$ mod L lies in W. Thus $\mathfrak{c}_p$ is a limit of bounding cycles, and $\mathfrak{h} = 0$.

If the group $\mathfrak{G}$ has the division-closure property (No. 8), we shall prove that the inverse of f is defined over the whole of $\mathfrak{H}'_p(\mathfrak{G})$. Suppose $\mathfrak{h}' = \{\mathfrak{h}^\alpha\}$ is an arbitrary element of $\mathfrak{H}'_p(\mathfrak{G})$. Let $\mathfrak{c}_p{}^\alpha$ be a representative of $\mathfrak{h}^\alpha$. Then $\mathfrak{c}_p{}^{\alpha+1} - \mathfrak{c}_p{}^\alpha$ reduced mod L^α is a limit of bounding cycles. As $\mathfrak{G}$ has the division-closure property, the group of bounding cycles of K mod L^α is closed. Hence there is a chain $\mathfrak{k}^\alpha{}_{p+1}$ such that

$$F(\mathfrak{k}^\alpha{}_{p+1}) = \mathfrak{c}_p{}^{\alpha+1} - \mathfrak{c}_p{}^\alpha \bmod L^\alpha.$$

Consider the sequence of chains

$$\mathfrak{c}_p{}^1,\ \mathfrak{c}_p{}^2 - F(\mathfrak{k}^1{}_{p+1}), \cdots, \mathfrak{c}_p{}^\alpha - \sum_{i=1}^{\alpha-1} F(\mathfrak{k}^i{}_{p+1}), \cdots$$

where $F(\mathfrak{k}^i{}_{p+1})$ is the boundary mod L of $\mathfrak{k}^i{}_{p+1}$. The $(\alpha + 1)$-st term is the α-th when reduced mod L^α. The sequence therefore converges to a cycle $\mathfrak{c}_p$ of

$\dot{K}$ mod L. Furthermore $\mathfrak{c}_p \sim \mathfrak{c}_p{}^a$ mod L^a; hence $\mathfrak{c}_p$ represents a class $\mathfrak{h} \in \mathfrak{H}_p(\mathfrak{G})$ such that $f(\mathfrak{h}) = \mathfrak{h}'$. It is not difficult to show that the inverse of f when defined is continuous.

Thus the two definitions of $\mathfrak{H}_p(K, L, \mathfrak{G})$ are equivalent if $\mathfrak{G}$ has the division-closure property. If we prefer the second definition then we can assert without qualification that $\mathfrak{X}$ is universal for the infinite cycles of an infinite complex.

It would be interesting to determine whether or not the two definitions are always equivalent.

Appendix I.

15. The homology groups of a bicompact connected group. We propose to establish the following

THEOREM 15. *If $\mathfrak{A}$ is a bicompact connected group, then $\mathfrak{A}$ and $\mathfrak{H}_1(\mathfrak{A}, \mathfrak{X})$ are bicontinuously isomorphic.*

The theorem is fairly trivial for a toral group. The general case is obtained by a limiting process to which the next few sections are devoted. We base our definition of homology groups on finite coverings by open sets. This we may do, for a topological group satisfies the separation axiom 5 of Hausdorff; so a bicompact group is normal (Remark I, No. 9). In Theorem 17, open coverings are used in an essential way, we have no proof of the theorem if the homology groups are based on closed coverings.

16. The induced homomorphism. Let A^1 and A^2 be topological spaces and f a continuous mapping of A^1 into a subset of A^2. Let $\{\phi^{i\alpha}\}$ $(i = 1, 2)$ be a complete system of finite coverings by open sets of A^i. Let $K^{i\alpha}$ be the nerve of $\phi^{i\alpha}$, and $\mathfrak{H}_p{}^{i\alpha} = \mathfrak{H}_p(K^{i\alpha}, \mathfrak{G})$ and $\mathfrak{H}_p{}^i = \mathfrak{H}_p(A^i, \mathfrak{G})$.

By means of f we shall define an inverse system S^{12} which includes $S^2 = \{\mathfrak{H}_p{}^{2\alpha}\}$ as a subsystem and $S^1 = \{\mathfrak{H}_p{}^{1\alpha}\}$ as a complete subsystem. S^{12} shall consist of the groups of the systems S^1 and S^2; it shall include all the homomorphisms of S^1 and S^2 and certain additional ones defined as follows. An open set in A^2 has an open set as its inverse image in A^1. Thus $\phi^{2\alpha}$ has as its inverse image a finite covering by open sets $\psi^{1\alpha}$ of A^1. Let $\phi^{1\beta}$ be a refinement of $\psi^{1\alpha}$. Then the projections $\phi^{1\beta} \to \psi^{1\alpha} \to \phi^{2\beta}$ determines a simplicial mapping of $K^{1\beta}$ into $K^{2\alpha}$. Let $H^{\alpha\beta}$ be the continuous homomorphism of $\mathfrak{H}_p{}^{1\beta}$ into $\mathfrak{H}_p{}^{2\alpha}$ induced by this simplicial mapping (No. 9). Let us include among the homomorphisms of S^{12} the homomorphisms $H^{\alpha\beta}$ for all α and related β satisfying the condition that $\phi^{1\beta}$ is a refinement of $\psi^{1\alpha}$. It is not difficult to prove that S^{12} is an inverse system.

As S^1 is a complete subsystem of S^{12}, $\mathfrak{H}_p{}^1$ is the limit group of S^{12}. The coördinates of an element $\mathfrak{h}^1$ of $\mathfrak{H}_p{}^1$ in the groups of S^2 make up the coördinates of an element $\mathfrak{h}^2$ of $\mathfrak{H}_p{}^2$. The correspondence $\tilde{f}(\mathfrak{h}^1) = \mathfrak{h}^2$ is a homomorphism of $\mathfrak{H}_p{}^1$ into $\mathfrak{H}_p{}^2$. It is continuous; for, if V^2 is a neighborhood of $\mathfrak{h}^2$ determined by a neighborhood V^{2a} of $\mathfrak{h}^{2a}$, then the neighborhood V^1 of $\mathfrak{h}^1$ determined by V^{2a} is such that $\tilde{f}(V^1) \subset V^2$.

The homomorphism $\tilde{f}$ of $\mathfrak{H}_p{}^1$ into $\mathfrak{H}_p{}^2$ is said to be induced by f.

We must show that $\tilde{f}$ is independent of the systems $\{\phi^{1a}\}$ and $\{\phi^{2a}\}$ used in its definition. But this is trivial; for new systems may be included with the old as complete subsystems of larger systems.

LEMMA 16. *If f and g are continuous mappings of A^1 into A^2 and A^2 into A^3, respectively, then the induced homomorphisms satisfy:* $(\widetilde{gf}) = \tilde{g}\tilde{f}$.

In S^{12} there are no homomorphisms of a group of S^2 into one of S^1. Similarly in S^{23}. Hence we may form the logical sum of S^{12} and S^{23} and obtain a new inverse system S^{123} which contains S^{13} as a complete subsystem. In S^{123} we may compare $(\widetilde{gf})$ and $\tilde{g}\tilde{f}$; and the assertion of the lemma follows.

17. On the homology groups of a limit space. Let $\{A^\alpha\}$ be an inverse system of bicompact spaces and A its limit space. We shall assume that, for all α, each point of A^α is the coördinate of an element of A. The set of groups $\{\mathfrak{H}_p{}^\alpha\}$ $(\mathfrak{H}_p{}^\alpha = \mathfrak{H}_p(A^\alpha, \mathfrak{G}))$ together with the homomorphisms induced by the mappings of the system constitutes, by Lemma 16, an inverse system. Let $\mathfrak{H}_p = \mathfrak{H}_p(A, \mathfrak{G})$. Then

THEOREM 17. *$\mathfrak{H}_p$ is the limit group of $\{\mathfrak{H}_p{}^\alpha\}$.*

Let $\{\phi^{(\alpha)\beta}\}$ be a complete system of finite coverings by open sets of A^α ((α) indicates that α is fixed). Each open set of $\phi^{\alpha\beta}$ has an open inverse image in A. Hence $\phi^{\alpha\beta}$ determines a finite covering $\psi^{\alpha\beta}$ of A by open sets. The double system $\{\{\psi^{\alpha\beta}\}\}$ is complete. For let ψ be an arbitrary finite covering of A by open sets. As A is bicompact, ψ has a refinement ψ' consisting of neighborhoods. Let the open set V^i $(i = 1, \cdots, k)$ of ψ' be defined by the neighborhood V^{α_i} in A^{α_i}. Let A^α be a common refinement of $A^{\alpha_1}, \cdots, A^{\alpha_k}$. Then the images $V^{\alpha i}$ of the V^i in A^α constitute a finite covering of A^α by open sets. Let $\phi^{\alpha\beta}$ be a refinement of this covering; then clearly $\psi^{\alpha\beta}$ is a refinement of ψ.

The nerve $K^{\alpha\beta}$ of $\phi^{\alpha\beta}$ is likewise the nerve of $\psi^{\alpha\beta}$, for the image of A in A^α covers A^α. Let $\mathfrak{H}_p{}^{\alpha\beta} = \mathfrak{H}_p(K^{\alpha\beta}, \mathfrak{G})$. Then $\mathfrak{H}_p$ is the limit group of the double system $\{\{\mathfrak{H}_p{}^{\alpha\beta}\}\}$. As $\psi^{\alpha\beta_1} \rightarrow \psi^{\alpha\beta_2}$ if and only if $\phi^{\alpha\beta_1} \rightarrow \phi^{\alpha\beta_2}$, $\mathfrak{H}_p{}^\alpha$ is the limit group of the subsystem $\{\mathfrak{H}_p{}^{(\alpha)\beta}\}$.

The coördinates in the subsystem $\{\mathfrak{H}_p^{(\alpha)\beta}\}$ of an element $\mathfrak{h} \in \mathfrak{H}_p$ make up an element $\mathfrak{h}^\alpha \in \mathfrak{H}_p^\alpha$. By the definition of the induced homomorphisms, the elements $\mathfrak{h}^\alpha$ are the coördinates of an element $\mathfrak{h}'$ of the limit group $\mathfrak{H}'_p$ of $\{\mathfrak{H}_p^\alpha\}$. The correspondence $f(\mathfrak{h}) = \mathfrak{h}'$ so defined is a bicontinuous isomorphism. It is trivial that f is a homomorphism. If $\mathfrak{h}'$ is an arbitrary element of $\mathfrak{H}'_p$, its coördinate $\mathfrak{h}^\alpha$ in $\mathfrak{H}_p^\alpha$ has a coördinate $\mathfrak{h}^{\alpha\beta}$ in $\mathfrak{H}_p^{\alpha\beta}$. Suppose $\mathfrak{H}_p^{\gamma\delta} \rightarrow \mathfrak{H}_p^{\alpha\beta}$. Let A^ϵ be a refinement of A^γ and A^α; and let $\psi^{\epsilon\lambda}$ be a refinement of $\psi^{\alpha\beta}$ and $\psi^{\gamma\delta}$. Then, by the definition of the induced homomorphism of $\mathfrak{H}_p^\epsilon$ into $\mathfrak{H}_p^\alpha$, we have $\mathfrak{h}^{\epsilon\lambda} \rightarrow \mathfrak{h}^{\alpha\beta}$. Similarly $\mathfrak{h}^{\epsilon\lambda} \rightarrow \mathfrak{h}^{\gamma\delta}$. From the uniqueness of projections $\mathfrak{h}^{\gamma\delta} \rightarrow \mathfrak{h}^{\alpha\beta}$. It follows that $\{\{\mathfrak{h}^{\alpha\beta}\}\}$ is an element $\mathfrak{h} \in \mathfrak{H}_p$ such that $f(\mathfrak{h}) = \mathfrak{h}'$. Therefore $f(\mathfrak{H}_p)$ covers $\mathfrak{H}'_p$.

If $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are distinct in $\mathfrak{H}_p$, then, for some pair α, β, $\mathfrak{h}_1^{\alpha\beta}$ and $\mathfrak{h}_2^{\alpha\beta}$ are distinct. Then $\mathfrak{h}_1^\alpha$ and $\mathfrak{h}_2^\alpha$ are distinct; and, finally $\mathfrak{h}'_1$ and $\mathfrak{h}'_2$ are distinct. Thus f is an isomorphism.

Let V' be a neighborhood of $\mathfrak{h}'$ in $\mathfrak{H}'_p$ determined by a neighborhood V^α of $\mathfrak{h}^\alpha$ in $\mathfrak{H}_p^\alpha$ which, in turn, is determined by a neighborhood $V^{\alpha\beta}$ of $\mathfrak{h}^{\alpha\beta}$ in $\mathfrak{H}_p^{\alpha\beta}$. Then $V^{\alpha\beta}$ determines a neighborhood V of $\mathfrak{h}$ in $\mathfrak{H}_p$ such that $f(V) \subset V'$. So f is continuous. On the other hand, if V is a neighborhood of $\mathfrak{h}$ in $\mathfrak{H}_p$ determined by a neighborhood $V^{\alpha\beta}$ of $\mathfrak{h}^{\alpha\beta}$ in $\mathfrak{H}_p^{\alpha\beta}$, $V^{\alpha\beta}$ determines a neighborhood V^α of $\mathfrak{h}^\alpha$ in $\mathfrak{H}_p^\alpha$, and this in turn determines a neighborhood V' of $\mathfrak{h}'$ in $\mathfrak{H}'_p$ satisfying $f(V) \supset V'$. So f is inversely continuous, and the theorem is proved.

18. Proof of Theorem 15. By Theorem 7.3, $\mathfrak{A}$ may be represented as the limit group of an inverse system $\{\mathfrak{A}^\alpha\}$ of elementary groups so that the image of $\mathfrak{A}$ in $\mathfrak{A}^\alpha$ covers the latter. As $\mathfrak{A}$ is connected, so is $\mathfrak{A}^\alpha$; hence $\mathfrak{A}^\alpha$ is a finite dimensional toral group. Let us express $\mathfrak{A}^\alpha$ as the direct sum $\sum_{i=1}^{k} \mathfrak{X}^{\alpha i}$ of groups isomorphic with $\mathfrak{X}$. The point set $\mathfrak{X}^{\alpha i}$ is a simple closed curve $\Gamma^{\alpha i}$ on $\mathfrak{A}^\alpha$. Giving to $\mathfrak{X}$ a definite orientation gives to each simple closed curve an orientation so that the set of 1-cycles (over $\mathfrak{J}$) so obtained form a 1-dimensional homology basis in $\mathfrak{A}^\alpha$.

If $\mathfrak{X}^\alpha$ is a homomorphic image of $\mathfrak{X}$ in $\mathfrak{A}^\alpha$, we may express $\mathfrak{X}^\alpha$ as a linear form $a_i\mathfrak{X}^{\alpha i}$ in the basis subgroups with integer coefficients (see No. 11). If Γ^α is the singular image on $\mathfrak{A}^\alpha$ of the basic 1-cycle of $\mathfrak{X}$, we assert that

$$\Gamma^\alpha \sim a_i\Gamma^{\alpha i}. \tag{18.1}$$

Let $\bar{\Gamma}^\alpha$ be the 1-cycle which is the singular image of the basic 1-cycle of $\mathfrak{X}$ under the mapping which sends $\mathfrak{x} \in \mathfrak{X}$ into $\mathfrak{x}\bar{\mathfrak{X}}^\alpha = \mathfrak{x}a_2\mathfrak{X}^{\alpha 2} + \cdots + \mathfrak{x}a_k\mathfrak{X}^{\alpha k}$. Let T be the 2-simplex in the (x, y)-plane having the vertices $(0, 0)$, $(0, 1)$, and $(1, 1)$.

Let C_1, C_2, C_3 be the edges $[(0,0),(0,1)]$, $[(0,1),(1,1)]$ $[(0,0),(1,1)]$ respectively. If T is suitably oriented, then $F(T) = C_1 + C_2 - C_3$. The transformation

$$f(x,y) = ya_1\mathfrak{X}^{\alpha 1} + x\bar{\mathfrak{X}}^{\alpha} \qquad (0 \leqq x \leqq 1;\ x \leqq y \leqq 1)$$

of T into $\mathfrak{A}^\alpha$ (on the right side x and y are real numbers mod 1) is continuous over $\bullet T$ and maps C_1 into $a_1\Gamma^{\alpha 1}$, C_2 into $\bar{\Gamma}^\alpha$, and C_3 into $\dot{\Gamma}^\alpha$. Thus $\Gamma^\alpha \sim a_1\Gamma^{\alpha 1} + \bar{\Gamma}^\alpha$. We may treat $\bar{\Gamma}^\alpha$ in the same way. Performing this operation k times, we obtain the relation (18.1).

We suppose that we have chosen basis subgroups $\mathfrak{X}^{\alpha i}$ $(i = 1, \cdots, k(\alpha))$ and corresponding 1-cycles $\Gamma^{\alpha i}$ on each group $\mathfrak{A}^\alpha$ of the system. If $\mathfrak{A}^\beta \to \mathfrak{A}^\alpha$, this homomorphism f may be written (No. 11)

(18.2) $$f(\mathfrak{x}\mathfrak{X}^{\beta i}) = \mathfrak{x}a_j{}^i\mathfrak{X}^{\alpha j}.$$

Then, by (18.1),

(18.3) $$f(\Gamma^{\beta i}) \sim a_j{}^i\Gamma^{\alpha j}.$$

Let $\mathfrak{H}^\alpha = \mathfrak{H}_1(\mathfrak{A}^\alpha, \mathfrak{X})$. $\mathfrak{H}^\alpha$ may be represented as the group of linear forms $\mathfrak{x}_i\Gamma^{\alpha i}$ where $\mathfrak{x}_i \in \mathfrak{X}$. Let I^α be the isomorphism which pairs $\mathfrak{x}_i\Gamma^{\alpha i}$ of $\mathfrak{H}^\alpha$ with $\mathfrak{x}_i\mathfrak{X}^{\alpha i}$ of $\mathfrak{A}^\alpha$.

If $\mathfrak{A}^\beta \to \mathfrak{A}^\alpha$, the induced homomorphism $\mathfrak{H}_p{}^\beta \to \mathfrak{H}_p{}^\alpha$, by (18.3), is

$$\bar{f}(\mathfrak{x}_i\Gamma^{\beta i}) = \mathfrak{x}_i a_j{}^i\Gamma^{\alpha j}.$$

It follows that the system of isomorphisms $\{I^\alpha\}$ establishes an isomorphism between the inverse systems $\{\mathfrak{A}^\alpha\}$ and $\{\mathfrak{H}^\alpha\}$. The limit group $\mathfrak{H}$ of $\{\mathfrak{H}^\alpha\}$ is therefore bicontinuously isomorphic with $\mathfrak{A}$. By Theorem 17, $\mathfrak{H}$ is $\mathfrak{H}_1(\mathfrak{A}, \mathfrak{X})$. This completes the proof.

Appendix II.

Example 1. We shall prove that the integers do not form a universal coefficient group for the homology theory of a compact space.

Let A^m $(m = 1, 2, \cdots)$ be a simple closed curve, and $M^m{}_{m+1}$ a continuous mapping of A^{m+1} into A^m of degree 2 (i. e. A^{m+1} is wrapped twice around A^m). Let A be the limit space of this inverse sequence.[14] Then A is 1-dimensional, bicompact, and it has the 2nd countability axiom. It may therefore be imbedded homeomorphically in euclidean 3-space. $\mathfrak{H}_1(A^m, \mathfrak{J})$ ($\mathfrak{J}$ = group of integers) is a free group on one generator. The induced homomorphism of $\mathfrak{H}_1(A^{m+1}, \mathfrak{J})$ into $\mathfrak{H}_1(A^m, \mathfrak{J})$ maps the generator of the first group into twice

[14] This example was considered by Vietoris 18.

the generator of the second. One proves readily that the limit group of $\{\mathfrak{H}_1(A^m, \mathfrak{J})\}$ reduces to the zero. By Theorem 17, this limit group is $\mathfrak{H}_1(A, \mathfrak{J})$. On the other hand, $\mathfrak{H}_1(A^m, \mathfrak{X})$ is isomorphic with $\mathfrak{X}$ and the homomorphism of $\mathfrak{H}_1(A^{m+1}, \mathfrak{X})$ into $\mathfrak{H}_1(A^m, \mathfrak{X})$ has the degree 2. It follows, by Lemma 2.1, that $\mathfrak{H}_1(A, \mathfrak{X})$ is not the zero group. In fact if $M^m{}_{m+1}$ is sufficiently smooth, it is easy to see that $\mathfrak{H}_1(A, \mathfrak{X})$ and A are homeomorphic. Since one cannot deduce the structure of $\mathfrak{H}_1(A, \mathfrak{X})$ from that of $\mathfrak{H}_1(A, \mathfrak{J})$ the statement is proved.

Example 2. We shall prove that $\mathfrak{X}$ is not a universal coefficient group for the homology theory of the finite cycles of an infinite complex.

Let A^m $(m = 1, 2, \cdots)$ be a simple closed curve and $M^m{}_{m+1}$ a continuous mapping of A^{m+1} into A^m of degree $m + 1$. Let A be the limit space of this sequence, and let A be imbedded homeomorphically in euclidean 3-space E_3. It is not difficult to prove that the 1-dimensional homology group over $\mathfrak{J}$ of the finite cycles of $E_3 - A$ is isomorphic with the group of rational numbers. Let A' be homeomorphic with A, and let it be imbedded in $E_3 - A$. Then the 1-dimensional homology group over $\mathfrak{J}$ of the finite cycles of $E_3 - (A + A')$ is the direct sum of two rational groups. Thus the two homology groups have ranks 1 and 2 respectively. However if we apply the theorem of Čech [9] to compute the 1-dimensional groups over $\mathfrak{X}$ of the finite cycles of $E_3 - A$ and $E_3 - (A + A')$ it is found that the first group is obtained by reducing $\mathfrak{X}$ modulo its subgroup of elements of finite order, and the second group is obtained by reducing a 2-dimensional toral group modulo its subgroup of elements of finite order. Both of these groups are isomorphic with the direct sum of groups of rational numbers equal in number to the power of the continuum. This proves the statement.

Example 3. We shall prove that, even if the coefficient group has the division-closure property, the group of bounding cycles in an infinite complex may not be closed. This is in contrast with the case of a finite complex (No. 8).

Let L' be the product complex of a circle by a line segment. We choose one of the two circles bounding L' and identify triplets of equally spaced points. The resulting complex consists of two 1-cycles Γ^1 and Γ^2 and a 2-chain L such that $F(L) = \Gamma^2 - 3\Gamma^1$. Let L_i $(i = 1, 2, \cdots)$ be a sequence of such complexes; and let us identify $\Gamma_i{}^1$ with $\Gamma^2{}_{i-1}$ $(i = 2, 3, \cdots)$ and give this cycle the new notation Γ_i. Let us add a 2-cell L_0 whose boundary is Γ_1. Then $F(L_0) = \Gamma_1$ and $F(L_i) = \Gamma_{i+1} - 3\Gamma_i$ $(i = 1, 2, \cdots)$ are the bounding relations in the resulting infinite complex K. As

$$\Gamma_i \sim 3\Gamma_{i-1} \sim 3^2\Gamma_{i-2} \sim \cdots \sim 3^{i-1}\Gamma_1 \sim 0,$$

the infinite 1-cycle $\sum_{i=1}^{\infty} \Gamma_i$ is a limit of the bounding cycles $\sum_{i=1}^{k} \Gamma_i$ $(k = 1, 2, \cdots)$. Suppose as is impossible, that there are integers x_i $(i = 0, 1, 2, \cdots)$ such that $F(\sum_{i=0}^{\infty} x_i L_i) = \sum_{i=1}^{\infty} \Gamma_i$. By a simple computation

$$F(\sum_{i=0}^{\infty} x_i L_i) = \sum_{i=1}^{\infty} (x_{i-1} - 3x_i)\Gamma_i.$$

It follows that $3x_i = x_{i-1} - 1$. Thus $| x_i | < | x_{i-1} | < \cdots < | x_0 |$. As the x's are integers, this cannot hold for every integer i. The contradiction proves the statement.

Example 4. We have proved in No. 10, that the p-th homology group of a topological space decomposes into a direct sum of a torsion group and a reduced homology group. In the special case of a finite complex, we have seen in No. 8 that the reduced homology group admits a further decomposition into the direct sum of two groups. We shall show by an example that this further decomposition does not occur in general in the homology groups of a compact metric space.

For a finite complex K, $\mathfrak{T}_p(K, \mathfrak{X})$ reduces to zero and $\mathfrak{H}_p(K, \mathfrak{X})$ is the direct sum of its component of zero and a finite group. We shall construct a 2-dimensional compact metric space A such that the component of zero of $\mathfrak{H}_2(A, \mathfrak{X})$ is not a direct summand of the entire group.

Let $\mathfrak{H}^*$ be the discrete group [15] generated by $e_1, e_2, \cdots$, and $a_0, a_1, a_2, \cdots$ subject to the relations $2^{2n}e_n = 0$ and $2a_n = a_{n-1} + e_n$ $(n = 1, 2, \cdots)$.

Let us prove that the subgroup $\mathfrak{E}$ of $\mathfrak{H}^*$ of elements of finite order is generated by $e_1, e_2, \cdots$. Suppose $a = \sum_1^n \alpha_i e_i + \sum_0^n \beta_i a_i$ is of finite order. Then $a' = \sum_0^n \beta_i a_i$ is of some finite order $k > 0$. It follows that there exist integers λ_i, μ_i $(i = 1, \cdots, m)$ giving the identity

$$k \sum_0^n \beta_i a_i \equiv \sum_1^m 2^{2i}\lambda_i e_i + \sum_1^m \mu_i(2a_i - a_{i-1} - e_i).$$

Comparing coefficients, we have

$$\begin{aligned} &\mu_i = 0 \ \ (i > n), \quad k\beta_n = 2\mu_n, \quad k\beta_i = 2\mu_i - \mu_{i+1} \ \ (i = 1, \cdots, n-1), \\ &k\beta_0 = -\mu_1, \quad 2^{2i}\lambda_i - \mu_i = 0 \qquad (i = 1, \cdots, m). \end{aligned}$$

We find that μ_1 is divisible by k; then, inductively, we find that μ_i

[15] For the construction of this group and the proof of its properties the author is indebted to Dr. Reinhold Baer.

$(i=1,\cdots,n)$ is divisible by k. Let $\mu_i/k=\tau_i$. Then $\beta_0=-\tau_1$ $\beta_i=2\tau_i-\tau_{i-1}$, $\beta_n=2\tau_n$, and

$$\sum_0^n \beta_i a_i = -\tau_1 a_0 + \sum_1^{n-1}(2\tau_i-\tau_{i+1})a_i + 2\tau_n a_n = \sum_1^n \tau_i(2a_i-a_{i-1}) = \sum_1^n \tau_i e_i.$$

Hence $g=\sum_1^n(\alpha_i+\tau_i)e_i$; as was to be proved.

We prove now that $\mathfrak{E}$ is not a direct summand of $\mathfrak{H}^*$. The difference group $\mathfrak{H}^*-\mathfrak{E}$ has $\bar{a}_0,\bar{a}_1,\cdots$ as generators with the relations $2\bar{a}_n=\bar{a}_{n-1}$ $(n=1,2,\cdots)$. If, to the contrary, $\mathfrak{E}$ is a direct summand, then there are elements f_i $(i=0,1,2,\cdots)$ in $\mathfrak{E}$ such that $2(a_i+f_i)=a_{i-1}+f_{i-1}$ $(i=1,2,\cdots)$. Using the relations in $\mathfrak{H}^*$, we find that

$$f_{i-1}=2f_i+e_i \qquad (i=1,2,\cdots).$$

Then, for each integer n,

$$f_0=2^nf_n+\sum_1^n 2^{i-1}e_i.$$

As f_n is an element of $\mathfrak{E}$ we can set $f_n=\sum_1^{m(n)}\alpha_i{}^n e_i$. We may assume without loss of generality that $m(n)\geqq n$. Then, if k is the order of f_0,

$$k2^n\sum_1^{m(n)}\alpha_i{}^n e_i + k\sum_1^n 2^{i-1}e_i = 0. \tag{1}$$

Let n be even and $i=n/2$. Then $k\equiv 0 \pmod{2^{n/2+1}}$. As this holds for every even integer n, we find that $k=0$ which contradicts the fact that k is the order of f_0 (if $f_0=0$, (1) holds for any integer k). This proves that $\mathfrak{E}$ is not a direct summand of $\mathfrak{H}^*$.

Let $\mathfrak{H}$ be the group of characters of $\mathfrak{H}^*$. The annihilator of $\mathfrak{E}$ in $\mathfrak{H}$ is the component of zero of $\mathfrak{H}$ (see [17], p. 386, Corollary 1c). As $\mathfrak{E}$ is not a direct summand of $\mathfrak{H}^*$, the component of zero of $\mathfrak{H}$ is not a direct summand of $\mathfrak{H}$ ([17], p. 382, Theorem 1b). We shall construct a compact metric space A such that $\mathfrak{H}=\mathfrak{H}_2(A,\mathfrak{X})$.

We first construct a direct sequence $\{\mathfrak{H}^{*m}\}$ whose limit group is $\mathfrak{H}^*$. $\mathfrak{H}^{*m}$ has $e_i{}^m$ $(i=1,\cdots,m)$ and c^m as its generators with the relations $2^{2i}e_i{}^m=0$ $(i=1,\cdots,m)$. The homomorphism H^* of $\mathfrak{H}^{*m}$ into $\mathfrak{H}^{*m+1}$ is defined by

$$H^*(c^m)=2c^{m+1}-e_{m+1}^{m+1}, \qquad H^*(e_i{}^m)=e_i{}^{m+1} \qquad (i=1,\cdots,m).$$

If we form the inverse sequence $\{\mathfrak{H}^m\}$ dual to $\{\mathfrak{H}^{*m}\}$ (No. 6), we find that

$$\mathfrak{H}^m=\mathfrak{X}^m+\sum_{i=1}^m \mathfrak{X}'^{mi}$$

where $\mathfrak{X}^m$ is isomorphic with $\mathfrak{X}$ and $\mathfrak{X}'^{mi}$ is isomorphic with the subgroup of $\mathfrak{X}$ of order 2^{2i}. The homomorphism H of $\mathfrak{H}_p{}^{m+1}$ into $\mathfrak{H}_p{}^m$ is given by

$$H(\mathfrak{x}\mathfrak{X}^{m+1}) = 2\mathfrak{x}\mathfrak{X}^m, \qquad H(\mathfrak{x}\mathfrak{X}'^{m+1,m+1}) = -\mathfrak{x}\mathfrak{X}^m,$$
$$H(\mathfrak{x}\mathfrak{X}'^{m+1,i}) = \mathfrak{x}\mathfrak{X}'^{mi} \qquad (i = 1, \cdots, m).$$

Let Q denote the point set in the plane consisting of the unit circle C and its interior. On C let us identify the point having the angular coördinate $\theta + p(2\pi/k)$ $(p = 1, \cdots, k;\ 0 \leqq \theta < 2\pi/k)$ with the point θ. In this way Q is converted into a complex Q^k such that $\mathfrak{H}_2(Q^k, \mathfrak{X})$ is isomorphic with the subgroup $\mathfrak{X}'^k$ of $\mathfrak{X}$ of order k. Let K^m $(m = 1, 2, \cdots)$ be a complex composed of a 2-sphere P^m and the complexes Q^{2i} $(i = 1, \cdots, m)$; we now denote these latter by $Q^{m,2i}$. Let π be a continuous mapping of K^{m+1} into K^m which maps P^{m+1} onto P^m with degree 2, $Q^{m+1,2i}$ onto $Q^{m,2i}$ $(i = 1, \cdots, m)$ with degree $+1$, and $Q^{m+1,2(m+1)}$ onto P^m with degree -1. It follows that $\mathfrak{H}_2(K^m, \mathfrak{X}) = \mathfrak{H}^m$ and the homomorphism of $\mathfrak{H}_2(K^{m+1}, \mathfrak{X})$ into $\mathfrak{H}_2(K^m, \mathfrak{X})$ induced by π is the homomorphism H of $\mathfrak{H}^{m+1}$ into $\mathfrak{H}^m$. If A is the limit space of the inverse sequence $\{K^m\}$, A is bicompact and has the 2nd countability axiom. As A possesses arbitrarily small mappings into 2-dimensional complexes, A is at most 2-dimensional. By Theorem 17, $\mathfrak{H}_2(A, \mathfrak{X})$ is the limit group of

$$\{\mathfrak{H}_2(K^m, \mathfrak{X})\} = \{\mathfrak{H}^m\}.$$

Thus $\mathfrak{H} = \mathfrak{H}_2(A, \mathfrak{X})$, and the component of zero of $\mathfrak{H}_2(A, \mathfrak{X})$ is not a direct summand. As $\mathfrak{H}_2(A, \mathfrak{X})$ is not zero, A is 2-dimensional.

Let A be imbedded in euclidean 5-space E_5. By the Pontrjagin theorem of duality [17], the 2-dimensional group over the integers of the finite cycles of $E_5 - A$ is the group $\mathfrak{H}^*$. In this the dual case the torsion group is not a direct summand of the homology group.

Example 5. We shall prove that $\mathfrak{X}$ considered as a discrete group is not a universal coefficient group. Thus the emphasis we have placed on the notion of a topologized homology group is essential.

Let $\mathfrak{A}$ be the character group of the discrete group of rational numbers. It is compact, connected, metric and 1-dimensional, and it contains no elements of finite order. If we construct in this group a Hamel basis, we find that $\mathfrak{A}$ is isomorphic with the direct sum of a set of groups each isomorphic with the group of rational numbers, the number of summands being the power of the continuum. The same is true of the direct sum $\mathfrak{A} + \mathfrak{A}$. Let $\mathfrak{X}_0$ denote the group $\mathfrak{X}$ with the discrete topology. As both $\mathfrak{X}_0$ and $\mathfrak{X}$ have the division-closure property, $\mathfrak{H}_1(\mathfrak{A}, \mathfrak{X})$ and $\mathfrak{H}_1(\mathfrak{A}, \mathfrak{X}_0)$ are isomorphic (though not continuously so). By Theorem 15, $\mathfrak{A}$ and $\mathfrak{H}_1(\mathfrak{A}, \mathfrak{X})$ are isomorphic, likewise $\mathfrak{A} + \mathfrak{A}$ and

$\mathfrak{H}_1(\mathfrak{A}+\mathfrak{A},\mathfrak{X})$. Then $\mathfrak{H}_1(\mathfrak{A},\mathfrak{X}_0)$ and $\mathfrak{H}_1(\mathfrak{A}+\mathfrak{A},\mathfrak{X}_0)$ are isomorphic. However, if $\mathfrak{R}$ is the group of rational numbers, it can be shown that $\mathfrak{H}_1(\mathfrak{A},\mathfrak{R})$ and $\mathfrak{H}_1(\mathfrak{A}+\mathfrak{A},\mathfrak{R})$ have ranks 1 and 2 respectively. So $\mathfrak{X}_0$ is not universal.

Bibliography.

1. J. W. Alexander, "On the chains of a complex and their duals," *Proceedings of the National Academy of Sciences*, vol. 21 (1935), pp. 509-511.
2. J. W. Alexander and L. W. Cohen, "A classification of the homology groups of compact spaces," *Annals of Mathematics*, vol. 33 (1932), pp. 538-566.
3. J. W. Alexander and L. Zippin, "Discrete abelian groups and their character groups," *Annals of Mathematics*, vol. 36 (1935), pp. 71-85.
4. P. Alexandroff, "Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension," *Annals of Mathematics*, vol. 30 (1929), pp. 101-187.
5. P. Alexandroff, "On local properties of closed sets," *Annals of Mathematics*, vol. 36 (1935), pp. 1-35.
6. P. Alexandroff and P. Urysohn, "Mémoire sur les espaces topologiques compacts," *Verhandelingen der Koninklijke Akad. te Amsterdam*, Eerste Sectie, vol. 14 (1929), pp. 1-96.
7. P. Alexandroff and H. Hopf, "Topologie," vol. 1, Berlin (1935).
8. E. Čech, "Théorie générale de l'homologie dans un espace quelconque," *Fundamenta Mathematik*, vol. 19 (1932), pp. 149-184.
9. E. Čech, "Les groupes de Betti d'un complexe infini," *Fundamenta Mathematik*, vol. 25 (1935), pp. 33-44.
10. E. Čech, "Concerning the Betti groups of a compact space," to appear in Menger's *Ergebnisse eines mathematische Kolloquiums* (1935).
11. H. Freudenthal, "Die R_n-adische Entwicklung von Räumen und gruppen," *Amsterdam Proceedings*, vol. 38 (1935), pp. 414-418.
12. F. Hausdorff, "Mengenlehre," 3rd edition, Berlin and Leipzig.
13. E. R. van Kampen, "Locally bicompact abelian groups and their character groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.
14. A. Kurosh, "Kombinatorischer Aufbau der bikompakten topologischen Räume," *Compositio Mathematische*, vol. 2 (1935), pp. 471-476.
15. S. Lefschetz, "Topology," *American Mathematical Society Publications*, vol. 12 (1930).
16. L. Pontrjagin, "Über den algebraischen Inhalt topologischer Dualitätssätze," *Mathematische Annalen*, vol. 105 (1931), pp. 165-205.
17. L. Pontrjagin, "The theory of topological commutative groups," *Annals of Mathematics*, vol. 35 (1934), pp. 361-390.
18. L. Vietoris, "Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen," *Mathematische Annalen*, vol. 97 (1927), pp. 454-472.

PRINCETON UNIVERSITY.

ON CLOSED SPACES OF CONSTANT MEAN CURVATURE.

By T. Y. Thomas.

Theorems on spaces of constant mean curvature (Einstein spaces) defined analytically by the equations

$$B_{\alpha\beta} = \lambda g_{\alpha\beta}, \qquad (B_{\alpha\beta} = \Sigma B^{\nu}{}_{\alpha\beta\nu}),$$

where λ is the mean curvature and the B's are the components of the curvature tensor, have been given by Kasner, Schouten and Struik.[1] These theorems are all of *local* character. In particular it has been shown that if the space is of dimensionality $n = 2, 3$ it must be of constant curvature. The following paper deals with *closed*[2] hypersurfaces S (without boundary) of constant mean curvature $\lambda > 0$ and dimensionality $n \geqq 2$ in a euclidean space E of $n + 1$ dimensions. In view of the above mentioned result, such spaces must be of constant curvature if $n = 2, 3$. We shall here prove that they must likewise be of constant curvature for $n \geqq 4$. As our proof depends on the fact that S is closed this result appears essentially as a theorem in the large and in this sense is distinguished from the local theorems of the above writers.

Our work is based on equations established in the paper by T. Y. Thomas, "On the variation of curvature in Riemann spaces of constant mean curvature," *Annali di Mathematica*, vol. 13 (1934-35), p. 227 and the paper by C. B. Allendoerfer, "Einstein spaces of class one," to appear in the *Bulletin of the American Mathematical Society*. From the latter of these we take the equations

$$(1) \qquad b_{\beta\gamma}b_{\alpha\delta} = k_1 g_{\beta\gamma}g_{\alpha\delta} + k_2 \sum_{a,b=1}^{n} g^{ab} \sum_{\mu,\nu=1}^{n} g^{\mu\nu}[2B_{a\alpha\nu\delta}B_{b\beta\gamma\mu} + B_{a\mu\delta\beta}B_{b\nu\gamma\alpha}],$$

where $$ds^2 = \sum_{a,b=1}^{n} g_{\alpha\beta}dx^{\alpha}dx^{\beta}, \qquad \psi = \sum_{a,b=1}^{n} b_{\alpha\beta}dx^{\alpha}dx^{\beta}$$

[1] E. Kasner, "The impossibility of Einstein fields immersed in flat space of five dimensions," *American Journal of Mathematics*, vol. 43 (1921), p. 126; "Finite representations of the solar gravitational field in flat space of six dimensions," *ibid.*, p. 130; "Geometrical theorems on Einstein's cosmological equations," *ibid.*, p. 217.

J. A. Schouten and D. J. Struik, "On some properties of general manifolds relating to Einstein's theory of gravitation," *American Journal of Mathematics*, vol. 43 (1921), p. 213.

[2] By saying briefly that the hypersurface S is closed, we mean in the terminology of the point set theory that S is compact and closed with respect to E. Hence S is contained in a finite portion of the euclidean space E.

are the first and second fundamental forms of the hypersurface S and

$$k_1 = \frac{\lambda}{n-2}, \qquad k_2 = \frac{1}{2\lambda(n-2)}.$$

From the former of these papers we select the equations

$$(2) \qquad \sum_{\mu,\nu=1}^{n} g^{\mu\nu} K_{,\mu\nu} - (4/3)\lambda K + \sum_{\mu,\nu=1}^{n} g^{\mu\nu} \sum_{\alpha,\beta,\gamma,\delta=1}^{n} B_{\alpha\beta\gamma\delta,\mu\nu} \lambda_1{}^{\alpha}\lambda_2{}^{\beta}\lambda_1{}^{\gamma}\lambda_2{}^{\delta} = 0.$$

Here $K_{,\mu\nu}$ are the components of the second extension of the sectional curvature K determined by the orthogonal unit vectors λ_1 and λ_2 at any point P of the hypersurface S and the $B_{\alpha\beta\gamma\delta,\mu\nu}$ are the components of the second extension of the curvature tensor B. We have the relations [3]

$$(3) \qquad \sum_{\mu,\nu=1}^{n} g^{\mu\nu} B_{\alpha\beta\gamma\delta,\mu\nu} = (2/3)\lambda B_{\alpha\beta\gamma\delta} + \sum_{a,b=1}^{n} g^{ab} \sum_{\mu,\nu=1}^{n} g^{\mu\nu} \times [B_{a\mu\alpha\beta} B_{b\nu\gamma\delta} + 2B_{\alpha\beta\delta\mu} B_{ba\gamma\nu} + 2B_{a\alpha\mu\delta} B_{b\beta\gamma\nu}].$$

Now interchange the indices γ, δ in (1) and subtract. When use is made of the Gauss equations relating the coefficients $b_{\alpha\beta}$ of the second fundamental form of the hypersurface S to the components of the curvature tensor we obtain

$$\begin{aligned}(4) \qquad B_{\alpha\beta\gamma\delta} &= k_1(g_{\beta\gamma}g_{\alpha\delta} - g_{\beta\delta}g_{\alpha\gamma}) \\ &+ k_2 \sum_{a,b=1}^{n} g^{ab} \sum_{\mu,\nu=1}^{n} g^{\mu\nu}[B_{a\mu\alpha\beta}B_{b\nu\gamma\delta} + 2B_{\alpha\beta\delta\mu}B_{ba\gamma\nu} + 2B_{a\alpha\mu\delta}B_{b\beta\gamma\nu}] \\ &+ k_2 \sum_{a,b=1}^{n} g^{ab} \sum_{\mu,\nu=1}^{n} g^{\mu\nu}[B_{a\mu\delta\beta}B_{b\nu\gamma\alpha} + B_{a\mu\gamma\beta}B_{b\nu\alpha\delta} + B_{a\mu\alpha\beta}B_{b\nu\delta\gamma}].\end{aligned}$$

We observe that when the last set of terms in (4) is multiplied by $\lambda_1{}^{\alpha}\lambda_2{}^{\beta}\lambda_1{}^{\gamma}\lambda_2{}^{\delta}$ and summed on repeated indices the expression vanishes identically. Hence in consequence of (2), (3), (4) and the equations defining the sectional curvature K [4] we have

$$\sum_{\mu,\nu=1}^{n} g^{\mu\nu} K_{,\mu\nu} = (2\lambda + (1/k_2))K - (k_1/k_2); \text{ or}$$

$$(5) \qquad \sum_{\mu,\nu=1}^{n} g^{\mu\nu} K_{,\mu\nu} = 2\lambda[(n-1)K - \lambda]$$

when we substitute the above values of the constants k_1 and k_2.

---

[3] These relations appear as equations (3.7) in Thomas (*loc. cit.*). Attention is here called to several typographical errors in the derivation of these relations. The term $B_{\alpha\beta\delta\epsilon,\gamma,\zeta}$ in the right member of (3.2) should be replaced by $B_{\alpha\beta\delta\epsilon,\gamma\zeta}$. Also in the equations at the bottom of p. 230 the term $B_{\alpha\beta\gamma\delta,\zeta,\epsilon}$ in the left member should be replaced by $B_{\alpha\beta\gamma\delta,\epsilon,\zeta}$ and the term $B_{\alpha\beta\gamma\delta,\epsilon,\zeta}$ in the right member should be replaced by $B_{\alpha\beta\gamma\delta,\epsilon\zeta}$.

[4] The equations (1.2) of Thomas (*loc. cit.*).

Since S is closed in the euclidean space E there exists a point P of S at which K assumes its maximum value. Since S is without boundary the point P will be an inner point of S and hence at P the left member of (5) will be $\leqq 0$. Suppose that S is not a space of constant curvature. Then at P the bracket expression will be > 0. Since $\lambda > 0$ by hypothesis it follows that the right member of (5) will be positive at the point P thus giving a contradiction. Hence the hypersurface S must have constant curvature.

We observe that the above requirements regarding continuity and differentiability are met if the hypersurface S is defined by functions $\phi^i(x)$ where $i = 1, \cdots, n + 1$ which are continuous and possess continuous partial derivatives to the fourth order.[5] Such a surface is said by some writers to be of class C^4. Using this terminology we have proved the following theorem.

THEOREM. *Any hypersurface S of class C^4 of constant mean curvature $\lambda > 0$ and dimensionality $n \geqq 4$ in a euclidean space E of $n + 1$ dimensions, the hypersurface S being closed but without boundary, is a space of constant curvature.*

In view of the above mentioned local theorems our theorem is proved if $n = 2, 3$ for hypersurfaces S of class C^3. We do not consider the question of whether the above theorem is valid for hypersurfaces of less restrictive class.

PRINCETON UNIVERSITY.

[5] In fact under this hypothesis the coefficients of the second fundamental form of S are continuous and possess continuous first and second derivatives. Cp. § 1 and § 2 of T. Y. Thomas, "Riemann spaces of class one and their characterization," to appear shortly in the *Acta Mathematica*. Hence it follows from the Gauss equations that the components of the curvature tensor are continuous with continuous first and second derivatives and this permits the definition of the second extension of the curvature tensor whose components occur in the equations (2) and (3).

ON CONTINUA OF CONDENSATION.

By G. T. WHYBURN.

A continuum M is said to have property N [1] provided that for every $\epsilon > 0$ there exists a finite collection G of disjoint non-degenerate subcontinua of M such that every subcontinuum of M of diameter $> \epsilon$ contains some continuum of G. Recently a study has been made by R. L. Moore [2] of this property in connection with various sorts of continua of condensation. Among other results, Moore proves that "The set of all regular curves with property N includes the set of all dendrons and all regular curves without continua of condensation and is included in the set of all regular curves with no essential continuum of condensation" (*loc. cit.*, p. 72).

In this paper I shall prove the following theorem [3] which yields exactly the relation between property N and the existence of continua of condensation in a given continuum:

THEOREM. *In order that a compact metric continuum M have property N it is necessary and sufficient that M be locally connected and that no cyclic element of M have a continuum of condensation.*

Proof. The condition is necessary. In the first place, since clearly if M has property N it cannot contain an infinite sequence of disjoint continua all of diameter greater than some $d > 0$, it follows that not only M but every subcontinuum of M must be locally connected. Thus the first condition is necessary.

Now let E be any true cyclic element of M, and suppose, contrary to our theorem, that E has a continuum of condensation K. Now obviously any subcontinuum of M must also have property N, so E has property N. Since, by the above, K must be locally connected, we can suppose without loss of generality that K is a simple arc ab. Let $\epsilon = \frac{1}{2}\rho(a, b)$. Since K is a continuum

[1] See R. L. Moore, "Fundamental point set theorems," *Rice Institute Pamphlet*, vol. 23, no. 1 (1936), see p. 67.

[2] *Loc. cit.* See also an abstract in *Bulletin of the American Mathematical Society*, vol. 42 (1936), p. 35.

[3] This theorem was recently communicated to R. L. Moore, who states that he had not thought of the theorem but that he can prove it with the aid of some results which he has found but not yet published.

of condensation of E, it follows by a theorem of the author's [4] that the non-local separating points of E are dense on K. Now let $W_1, W_2, \cdots, W_n$ be a set of disjoint continua such that every subcontinuum of E of diameter $> \epsilon$ contains a continuum W_i. Let $W_{n_1}, W_{n_2}, \cdots, W_{n_j}$ be the continua W_i which lie wholly in K and let Y be the sum of the remaining ones. Then for each $i \leq j$, W_{n_i} is a subarc $a_i b_i$ of ab; hence W_{n_i} contains an inner point which is not a local separating point of E, and accordingly there exists an arc $\overline{x_i y_i}$ in E such that x_i, y_i are in the order a, a_i, x_i, y_i, b_i, b, $\overline{x_i y_i} \cdot K = x_i + y_i$, and $\overline{x_i y_i} \cdot Y = 0$. Now for each $i \leq j$, replace the arc $x_i y_i$ of ab by the arc $\overline{x_i y_i}$ and call the continuum thus formed H. Then since $H \supset a + b$, we have $\delta(H) > \epsilon$. But since for no $i \leq j$ does H contain the arc $x_i y_i$ of K, H contains no set W_{n_i}; and since $\overline{x_i y_i} \cdot Y = 0$ for each $i \leq j$, H can contain no set of Y. Hence H contains no set W_i whatever. Thus the supposition that the second condition is not necessary leads to a contradiction.

The proof for the sufficiency of the conditions will be given in five steps.

(1) *Any continuum M having no continuum of condensation has property N.*[5]

Proof. Let $\epsilon > 0$ be given. We can write [6]

$$M = F + \sum_{i=1,2,\ldots} \widehat{a_i x_i b_i} \tag{i}$$

where F is closed and totally disconnected and each set $\widehat{a_i x_i b_i}$ is an open free arc. Since at most a finite number of the arcs $\widehat{a_i x_i b_i}$ are of diameter $\geq \epsilon/5$, by adding a finite number of points on these to F clearly we can obtain the decomposition (i) so that all arcs $\widehat{a_i x_i b_i}$ are of diameter $< \epsilon/5$. Let us suppose this has been done.

Now since F is totally disconnected and closed, it follows at once that there exists an integer k such that every component of $M - \sum_{i=1}^{k} \widehat{a_i x_i b_i}$ is of diameter $< \epsilon/5$. For otherwise we could find a monotone decreasing sequence of continua $K_1, K_2, K_3, \cdots$ such that for each k, $\delta(K_k) \geq \epsilon/5$ and K_k is a component of $M - \sum_{i=1}^{k} \widehat{a_i x_i b_i}$; this is impossible since then ΠK_i would be a continuum in F of diameter $\geq \epsilon/5$.

For each $i \leq k$, let W_i be any closed arc contained wholly in $\widehat{a_i x_i b_i}$. Then

[4] See *Mathematische Annalen*, vol. 102 (1929), p. 320.

[5] See Moore, *loc. cit.* The proof is given here merely for the sake of completeness.

[6] See Urysohn, *Verhandelingen der Akademie te Amsterdam*, vol. 13, no. 3 (1927), p. 57.

every subcontinuum of M of diameter $> \epsilon$ must contain one of these arcs W_i. For let Q be a subcontinuum of M of diameter $> \epsilon$. Then Q intersects at least two components of $M - \sum_{i=1}^{k} \widehat{a_ix_ib_i}$ since if L is any one such component, then $L +$ all arcs $\widehat{a_ix_ib_i}$ having an endpoint on L is a set of diameter $< 3\epsilon/5$. Let $a, b \in Q$, where a and b lie in different components C_a and C_b of $M - \sum_{i=1}^{k} \widehat{a_ix_ib_i}$. Let ab be an arc in Q. Let a' be the last point of C_a on ab in the order a, b and let b' be the first point following a' which belongs to $M - \sum_{i=1}^{k} \widehat{a_ix_ib_i}$. Then clearly $\widehat{a'b'}$ is one of the arcs $\widehat{a_ix_ib_i}$ for some $i \leq k$. Whence $Q \supset a'b' \supset W_i$.

(2) *If no true cyclic element of a locally connected continuum S has a continuum of condensation, neither does any cyclic chain of S.*

For let Q be a subcontinuum of a cyclic chain $C(a, b)$. Then Q contains an arc pq, and pq has a segment xy which either belongs to the set K of all points separating a and b in $C(a, b)$ or to a true cyclic element C_i in $C(a, b)$. If $xy \subset K$, clearly $xy - (x + y)$ is open in $C(a, b)$. If $xy \subset C_i$, then $[xy - (K \cdot xy)] \cdot \overline{C(a, b) - C_i} = 0$; and since xy is not a continuum of condensation of C_i, some subarc $x'y'$ of xy is such that $x'y' \cdot \overline{C(a, b) - Q} = 0$. Thus in either case Q is not a continuum of condensation of $C(a, b)$.

(3) *If no true cyclic element of a locally connected continuum S has a continuum of condensation and if we express*

$$S = \sum_{i=1}^{k} C(p_i, q_i) + \sum_{i=k+1}^{\infty} C(p_i, q_i) + H$$

as in the cyclic chain approximation theorem,[7] *then for no k does $H_k = \sum_{i=1}^{k} C(p_i, q_i)$ have a continuum of condensation. Thus every H_k has property N.*

(4) *If every H_k in a locally connected continuum S has property N, so also does S.*

Proof. Let $\epsilon > 0$. Then there exists a k such that every component of $S - H_k$ is of diameter $< \epsilon/3$. Since H_k has property N, there exists a finite

[7] See Kuratowski and Whyburn, *Fundamenta Mathematicae*, t. 16 (1930), pp. 305-331. Here $C(p_i, q_i)$ is a cyclic chain.

number of disjoint continua $W_1, W_2, \cdots, W_n$ in H_k such that any subcontinuum of H_k of diameter $> \epsilon/3$ contains one of these continua W_i.

Now let Q be an arbitrary subcontinuum of S of diameter $> \epsilon$. Let $p, q \varepsilon Q$ be chosen so that $\rho(p, q) > \epsilon$. Let $p' = p$ if $p \varepsilon H_k$ and if not let p' be the boundary of the component Q_p of $S - H_k$ containing p. Similarly let $q' = q$ if $q \varepsilon H_k$ and otherwise let q' be the boundary of the component Q_q of $S - H_k$ containing q. Then since

$$\rho(p, p') + \rho(q, q') \leq \delta(Q_p) + \delta(Q_q) < \epsilon/3 + \epsilon/3 = 2\epsilon/3,$$

we have

$$\rho(p', q') > \epsilon/3.$$

Thus since $p' + q' \subset Q \cdot H_k$, we have

$$\delta(Q \cdot H_k) > \epsilon/3.$$

But $Q \cdot H_k$ is a continuum. Whence $Q \cdot H_k \supset W_j$ for some j. Accordingly, S has property N.

UNIVERSITY OF VIRGINIA.

CONTINUOUS TRANSFORMATIONS PRESERVING ALL TOPOLOGICAL PROPERTIES.

By JAMES F. WARDWELL.

1. *Introduction.* This paper concerns itself with a solution to the following problem:[1] If A and B are any compact metric spaces and $T(A) = B$ is a continuous transformation, under what conditions will B be homeomorphic with A, that is, what continuous transformations will preserve all topological properties of A? In view of the fact that continuous transformations and upper semi-continuous decompositions[2] are known to be equivalent[3] for compact metric spaces, the above problem can be stated in this way: If we have an upper semi-continuous decomposition of a compact metric space A, under what conditions will the hyperspace B of this decomposition be the same kind of space as A, that is, be homeomorphic with A?

The solution to this problem, for the case when A is a plane or sphere, is due to R. L. Moore.[4] It may be stated as follows: *If A is a topological sphere and $T(A) = B$ is a monotone transformation*[5] *such that, for any $b \varepsilon B$, $T^{-1}(b)$ does not separate A, then B is a topological sphere.*

However, no conditions have heretofore been found which yield the desired result for general compact metric spaces or even for any compact Euclidean spaces of higher dimension than two.

2. *Conditions.* A little investigation makes it clear that the conditions which are to be imposed must be conditions on the complements of the sets of the decomposition (that is, the sets $T^{-1}(b)$, for $b \varepsilon B$) in the space as well as

[1] This problem was suggested by G. T. Whyburn to whom the author is greatly indebted for his helpful suggestions and criticism in the preparation of this paper. See the abstract of his paper "Analytic topology" in *American Mathematical Monthly*, vol. 42 (1935), p. 190.

[2] See R. L. Moore, *Transactions of the American Mathematical Society*, vol. 27 (1925), pp. 416-428.

[3] See P. Alexandroff, *Mathematische Annalen*, vol. 96 (1927), pp. 551-571, and C. Kuratowski, *Fundamenta Mathematicae*, vol. 11 (1928), pp. 169-185.

[4] See *loc. cit.*

[5] A continuous transformation $T(A) = B$ is said to be a *monotone transformation* when each set $T^{-1}(b)$, for $b \varepsilon B$, is connected. See C. B. Morrey, Jr., *American Journal of Mathematics*, vol. 57 (1935), pp. 17-50, and G. T. Whyburn, *American Journal of Mathematics*, vol. 56 (1934), no. 2, pp. 294-302.

on these sets themselves. It seemed that the following condition might yield the desired result:

I. *For any* $b \varepsilon B$ *and any* $x \varepsilon A$, *there exists a homeomorphism* $W(A - x) = A - T^{-1}(b)$.

However we have not been able to show that, for arbitrary compact metric spaces A and B, B is homeomorphic with A when this condition is satisfied, even when almost all of the sets $T^{-1}(b)$ are degenerate. The same was true for this condition:

II. *For any* $b \varepsilon B$ *and any* $x \varepsilon T^{-1}(b)$, *there exists a homeomorphism* $W(A - x) = A - T^{-1}(b)$.

This last condition is less restrictive than the first one in that it does not make the space A homogeneous while condition I does. If II is satisfied and if A is homogeneous, then I is obviously satisfied.

We finally found that if condition II was further restricted to give:

III. *For any* $\epsilon > 0$, *any* $b \varepsilon B$, *and any* $x \varepsilon T^{-1}(b)$, *there exists a homeomorphism* $W(A - x) = A - T^{-1}(b)$ *which is stationary*[6] *outside of the* ϵ*-neighborhood of* $T^{-1}(b)$.

the desired results are obtained for the case stated in section 7 below.

3. *Some effects of these conditions.*

If the compact metric space A *is connected, conditions I and III each imply that* $T(A) = B$ *is a monotone transformation.*

In order to demonstrate this result, let us assume the contrary in both cases. Then there exists some set $T^{-1}(b) = C$, where $C = C_1 + C_2$, mutually separated. Now $\bar{C}_1 \cdot \bar{C}_2 = 0$, since C is closed. Take neighborhoods U_1 and U_2 of C_1 and C_2 respectively so that $\bar{U}_1 \cdot \bar{U}_2 = 0$. Let $U = U_1 + U_2$. There exists a closed cutting S in $A - U$ which separates $\bar{U}_1$ and $\bar{U}_2$ in A, that is, $A - S = A_1 + A_2$, mutually separated, where

$$A_1 \supset \bar{U}_1 \supset C_1 \quad \text{and} \quad A_2 \supset \bar{U}_2 \supset C_2.$$

For condition I, take any $x \varepsilon A - S$. Now there exists a homeomorphism $W(A - x) = A - C$. If we apply W^{-1} to the set

$$A - C - S = (A_1 - C_1) + (A_2 - C_2)$$

[6] A transformation W is said to be *stationary* over all points y for which $W(y) = y$.

we obtain a separation

$$A - x - W^{-1}(S) = W^{-1}(A_1 - C_1) + W^{-1}(A_2 - C_2).$$

Now $x + W^{-1}(S)$ is a closed cutting of A, and x is an isolated point of this cutting. Furthermore x is a limit point of both $W^{-1}(A_1 - C_1)$ and $W^{-1}(A_2 - C_2)$. Therefore, by an established theorem,[7] x is a local separating point[7] of A. Now for any point y in $A - x$, there exists a homeomorphism $R(A - x) = A - y$. Hence every point of A is a local separating point of A. Now A is homogenous, and hence every point of A is of the same Menger order. Accordingly no point of A is of order greater than 2 because, by the Local Separating Point–Order Theorem,[7] there exists at most a countable number of local separating points of order greater than 2. Since A consists of more than one point and is connected, it follows that every point of A is of order exactly 2. Therefore A is a simple closed curve. From this it follows that $C_1 + C_2$ separates A; and thus x separates A, which is a contradiction, since no single point separates a simple closed curve. Hence, for condition I, every set $T^{-1}(b)$ is connected, that is, T is a monotone transformation.

For condition III, take any point $x \varepsilon C_1$. By hypothesis, there exists a homeomorphism $W(A - x) = A - C$ which is stationary outside of U. Let $R(A) = A$ be the transformation such that, for any $y \varepsilon A - C$, $R(y) = W^{-1}(y)$ while $R(C) = x$. Now R is continuous. Let the cutting $S = A - U$. Then we have $A - S = A - (A - U) = U_1 + U_2$, mutually separated. Applying R to $A - S$ we have: $A - R(S) = R(U_1) + R(U_2)$. However, $R(S) = W^{-1}(S) = S$, since W is stationary in $A - U$. Therefore, $A - S = R(U_1) + R(U_2)$. Let P be the set of all points p of U_2 so that $R(p) \subset U_1$. Consequently $R(U_2 - P) \subset U_2$. If we let $Q = S + U_2 - P$, then $A = Q + U_1 + P$. Now it can easily be seen that $Q + U_1$ and P are mutually separated sets. Hence we have a separation of the connected set A; and this is a contradiction. Therefore T is a monotone transformation also in this case.

Since conditions I and III each make the transformation T a monotone one when A is connected, it follows that each of these conditions reduces to those of Moore for the case of the plane.

Condition II does not make T a monotone transformation even when A is connected.

This result is illustrated by the following example. Take a line L in a plane and a segment pq on this line. In the interior of pq take a point p_1,

[7] See G. T. Whyburn, *Monatshefte für Mathematik und Physik*, vol. 36 (1929), pp. 305-314.

and in the interior of p_1q take a point q_1. In pp_1 take a sequence $\{p_i\}$ of points converging to p, ordering these points so that $p_{i+1} \subset pp_i$, for each i. Take a similar sequence $\{q_i\}$ of points in q_1q converging to q. On one side of the line L construct some simple arcs as follows: Take a simple arc C from p_1 to q_1 which meets pq only in the points p_1 and q_1. For each i, take a simple arc C_i from p_i to p_{i+1} which meets the set $pq + C + \sum_{j=1}^{i-1} C_j$ only in p_i and p_{i+1}, and such that $\delta(C_i) \to 0$. For the points of $\{q_i\}$, for each i, take a simple arc D_i from q_i to q_{i+1} which meets $pq + C + \sum_{k=1}^{\infty} C_k + \sum_{j=1}^{i-1} D_j$ only in the points q_i and q_{i+1}, and such that $\delta(D_i) \to 0$. On the other side of L construct sequences of simple closed curves as follows: At each point p_i take a sequence $\{E_k^i\}$ of simple closed curves converging to p_i, every two of which intersect only in p_i, and so that $E_k^i \cdot E_m^j = 0$, for $i \neq j$, and for all k and m. At each point q_i take a sequence $\{F_k^i\}$ of simple closed curves converging to q_i, every two of which intersect only in q_i, and so that $F_k^i \cdot F_m^j = 0$, for $i \neq j$, and for all k and all m, and no one of which intersects any of the E_k^i, for all i and all k. Let A represent the set

$$pq + C + \sum_{i=1}^{\infty} (C_i + D_i) + \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} (E_k^i + F_k^i).$$

Now A is a connected compact metric space. Take a decomposition of A into the set $p_1 + q_1$ and the points of $A - (p_1 + q_1)$. This is obviously an upper semi-continuous decomposition. We must now show that $A - (p_1 + q_1)$ is homeomorphic with $A - p_1$ and with $A - q_1$. Now $A - (p_1 + q_1)$ consists of the sets:

$$(pp_1 + \sum_{i=1}^{\infty} C_i + \sum_{i=2}^{\infty}\sum_{k=1}^{\infty} E_k^i) - p_1,$$
$$(q_1q + \sum_{i=1}^{\infty} D_i + \sum_{i=2}^{\infty}\sum_{k=1}^{\infty} F_k^i) - q_1,$$

and the countable collection of free open arcs:

$$C - (p_1 + q_1),\ p_1q_1 - (p_1 + q_1),\ E_k^1 - p_1, \text{ and } F_k^1 - q_1, \text{ for all } k.$$

Furthermore, $A - p_1$ and $A - q_1$ each consist of the same number and same types of sets as $A - (p_1 + q_1)$, and it is easily seen that both $A - p_1$ and $A - q_1$ are homeomorphic with $A - (p_1 + q_1)$.

If A is a 2-dimensional manifold, conditions I and III are equivalent on A.

For, if condition III is given, condition I is satisfied because A is homo-

geneous. If condition I is given, we have, for any set $T^{-1}(b)$, for $b \varepsilon B = T(A)$, and any $x \varepsilon A$, a homeomorphism $W(A - x) = A - T^{-1}(b)$. Now take the point x in $T^{-1}(b)$. The set $T^{-1}(b)$ is connected since A is connected. Take a monotone sequence $\{U_i\}$ of neighborhoods closing down on x such that $\bar{U}_i$ is a closed 2-cell, for each i. Let $W(U_i - x) = V_i^\circ$ and $V_i^\circ + T^{-1}(b) = V_i$, for each i. The neighborhoods V_i close down on $T^{-1}(b)$ since the neighborhoods U_i close down on x. Let $C_i = \bar{U}_i - \bar{U}_{i+1}$ and $D_i = V_i - \bar{V}_{i+1}$, for each i. Now $W[F(U_i^\circ)] = F(V_i)$,[8] for each i, and W is a homeomorphism. Hence for each i, $C_i + F(U_i) + F(U_{i+1})$ and $D_i + F(V_i) + F(V_{i+1})$ are each homeomorphic with a circular ring. For any $\epsilon > 0$ take an ϵ-neighborhood U_ϵ of $T^{-1}(b)$ so that $U_\epsilon - T^{-1}(b)$ is homeomorphic with a 2-cell V_ϵ of x minus x. Pick a k large enough so that $\bar{V}_k \subset U_\epsilon$ and $\bar{U}_k \subset U_\epsilon$. Now there exists a homeomorphism

$$R_\epsilon[(U_\epsilon - \bar{U}_k) + F(U_k) + F(U_\epsilon)] = (U_\epsilon - \bar{V}_k) + F(V_k) + F(U_\epsilon),$$

where $R_\epsilon(p) = p$ for $p \varepsilon F(U_\epsilon)$, since each of these sets is a circular ring and clearly any homeomorphism between the two outer curves of two such rings can be extended to the whole rings.[9] Let $R_\epsilon = S_{k-1}$. Similarly there exists a homeomorphism

$$S_k[C_k + F(U_k) + F(U_{k+1})] = D_k + F(V_k) + F(V_{k+1})$$

such that $S_k(x) = S_{k-1}(x)$ for $x \varepsilon F(U_k)$. Likewise there exists a homeomorphism

$$S_{k+1}[C_{k+1} + F(U_{k+1}) + F(U_{k+2})] = D_{k+1} + F(V_{k+1}) + F(V_{k+2})$$

such that $S_{k+1}(x) = S_k(x)$ for $x \varepsilon F(U_{k+1})$, and so on. In general for $j \geq -1$ we have a homeomorphism

$$S_{k+j}[C_{k+j} + F(U_{k+j}) + F(U_{k+j+1})] = D_{k+j} + F(V_{k+j}) + F(V_{k+j+1})$$

such that $S_{k+j}(x) = S_{k+j-1}(x)$ for $x \varepsilon F(U_{k+j})$.

Now define a transformation S as follows: $S(p) = p$ if $p \varepsilon M - U_\epsilon$, $S(p) = R_\epsilon(p)$ if $p \varepsilon (U_\epsilon - \bar{U}_k) + F(U_k) + F(U_\epsilon)$, $S(p) = S_j(p)$ for $j \geq k$ where j is the least integer such that p is a point of $C_j + F(U_j) + F(U_{j+1})$. Now $S(A - x) = A - T^{-1}(b)$ is a homeomorphism which is stationary outside of U_ϵ. Hence condition III is satisfied.

[8] The boundary $\bar{P} - P$ of any open set P is represented by $F(P)$.

[9] See Schoenflies, *Mathematische Annalen*, vol. 62 (1906), p. 324.

4. *Notation.* If A and B are any compact metric spaces and $T(A) = B$ is a continuous transformation, let $G = G_0$ denote the collection of all non-degenerate sets $T^{-1}(b)$, for $b \varepsilon B$ in A. Let G_1 denote the collection of all sets of G which intersect $L = L_0 = \text{lim. sup. } G$. Let G_2 be the collection of all sets of G_1 which intersect $L_1 = \text{lim. sup. } G_1$, etc.

If the collection G is countable, let $\{g_i^k\}$ represent the sets of $G_{k-1} - G_k$, for $k = 1, 2, \cdots$. Cover each g_i^k by a neighborhood U_i^k so that: (a) $U_i^k \cdot g_j^k = 0$, for $j \neq i$; (b) $U_i^k \cdot \sum_{j=1}^{i-1} \bar{U}_j^k = 0$; (c) no set of G_k intersects U_i^k, for any i; and (d) $\bar{U}_i^k$ is contained in the $\rho(g_i^k, L_{k-1})/2$-neighborhood of g_i^k, for each i.

5. LEMMA. *If A and B are compact metric spaces and $T(A) = B$ is continuous and such that for any $\epsilon > 0$, any $b \varepsilon B$, and any $x \varepsilon T^{-1}(b)$, there exists a homeomorphism $W(A - x) = A - T^{-1}(b)$ which is stationary outside of the ϵ-neighborhood of $T^{-1}(b)$, and if there exists an integer k so that $G_k = 0$, then A is homeomorphic with B.*

Proof. 1°. If $G_i = 0$, for any integer i, then G_{i-1} is a null sequence.[10] For, if not, there exists a number $d > 0$ such that there is an infinite collection $\{h_i\}$ of sets of G_{i-1} the diameter $\delta(h_i)$ of each of which is greater than d. Hence, if we take a convergent subsequence $\{h'_i\}$ of $\{h_i\}$, $\delta(\text{lim. sup. } \{h'_i\}) \geq d$. Now lim. sup. $\{h'_i\} \subset T^{-1}(p)$ for some $p \varepsilon B$. Therefore $\delta(T^{-1}(p)) \geq d$, and thus $T^{-1}(p) \varepsilon G_i$, since L_{i-1} contains lim. sup. $\{h_i\}$. This contradicts the fact that $G_i = 0$.

We shall make the proof by induction.

2°. We shall first demonstrate the result for the case $k = 1$. From 1° it follows that G is a null sequence, since $G_1 = 0$. Now $\{g_i^1\}$ denotes the collection of sets of G. Take a point $y_i \varepsilon g_i^1$, for each i. Let $T(g_i^1) = b_i$, for each i. By hypothesis, there exists a homeomorphism[11] $W_1(A - y_1) = A - g_1^1$ which is stationary in $A - U_1^1$. Now $T W_1(A - y_1) = B - b_1$. Let $T_1(A) = B$ be the transformation such that for any $x \varepsilon A$, $T_1(x) = T W_1(x)$ when $x \varepsilon A - y_1$, and $T_1(y_1) = b_1$. Now T_1 is univalued and continuous, since T and W_1 are, and since A is compact. Now T_1 is such that $T_1^{-1}(b_i) = g_i^1$, for $i = 2, \cdots$, since $T_1 = T$ over $A - U_1^1$. Similarly, for

---

[10] A sequence $\{M_i\}$ of sets is said to be a *null sequence* provided that, for any $\epsilon > 0$, there are at most a finite number of the sets which have a diameter greater than ϵ.

[11] The method used here was suggested by a proof given by Mrs. Lucille Whyburn for a certain finite case of our problem.

each n, there exists a homeomorphism $W_n(A - y_n) = A - g_n^1$ which is stationary in $A - U_n^1$. Moreover $g_n^1 \equiv T^{-1}_{n-1}(b_n)$, and thus

$$W_n(A - y_n) = A - T^{-1}_{n-1}(b_n).$$

Hence $T_{n-1}W_n(A - y_n) = B - b_n$. Let $T_n(A) = B$ be the transformation such that, for $x \varepsilon A - y_n$, $T_n(x) = T_{n-1}W_n(x)$, and $T_n(y_n) = b_n$. Now T_n is also univalued and continuous. Furthermore $T_n^{-1}(b_i) = g_i^1$, for $i > n$, and $T_n^{-1}(b_j)$ is a single point for each $j \leq n$. From our definition of $T_n(A) = B$ it follows that, for any n, and any $x \varepsilon A$:

$$\begin{array}{ll}
T_n(x) = T_{n-1}W_n(x) & \text{except for } x = y_n, \\
T_n(x) = T_{n-2}W_{n-1}W_n(x) & \text{except for } x = y_n \text{ or } y_{n-1}, \\
\cdot \quad\quad \cdot & \cdot \quad \cdot \quad \cdot \\
\cdot \quad\quad \cdot & \cdot \quad \cdot \quad \cdot \\
\cdot \quad\quad \cdot & \cdot \quad \cdot \quad \cdot \\
\cdot \quad\quad \cdot & \cdot \quad \cdot \quad \cdot \\
T_n(x) = T_1W_2 \cdots W_n(x) & \text{except for } x \varepsilon \sum_{i=2}^{n} y_i, \\
T_n(x) = T\, W_1W_2 \cdots W_n(x) & \text{except for } x \varepsilon \sum_{i=1}^{n} y_i.
\end{array}$$

Let $S = \lim_{n=\infty} \{T_n\}$. We will now prove that $\{T_n\}$ is uniformly convergent. Since all T_n are continuous over A, it will follow that S is also continuous. Since T is continuous and is defined over a compact space, it is uniformly continuous. Hence, for any $\epsilon > 0$, there exists a δ_ϵ such that, if $\rho(x, y) < \delta_\epsilon$, for $x, y \varepsilon A$, $\rho(T(x), T(y)) < \epsilon$. Take any $\epsilon > 0$. Let $\delta = \delta_{\epsilon/2}$. Cover each point of L by a $\delta/2$-neighborhood $V_{\delta/2}$ and by a δ-neighborhood V_δ. Since L is closed and compact we can find a finite number of the neighborhoods $V_{\delta/2}$, say $V^1{}_{\delta/2}, V^2{}_{\delta/2}, \cdots V^h{}_{\delta/2}$, whose sum covers L. Let $V_\delta^1, V_\delta^2, \cdots V_\delta^h$, be the corresponding δ-neighborhoods. Now $V = \sum_{i-1}^{h} V_\delta^i \supset L$. Since $\{g_i^1\}$ is a null sequence and because of our definition of the neighborhoods U_i^1, it follows that $\{U_i^1\}$ s a null sequence and lim. sup. $\{U_i^1\} = L$. Hence there exists an integer N such that $\delta(U_n^1) < \delta/2$, for all $n > N$. There also exists an integer M so that $U_m \cdot \sum_{i=1}^{h} V^i{}_{\delta/2} \neq 0$, for all $m > M$. For, if not, there would be an infinite collection of the sets of $\{U_i^1\}$ which did not intersect $\sum_{i=1}^{h} V^i{}_{\delta/2}$. Since A is compact, this collection would possess a limit point p which is contained in L but not contained in $\sum_{i=1}^{h} V^i{}_{\delta/2}$, which is a contradiction.

Let N_ϵ be the larger of the two numbers N and M. For any $n > N_\epsilon$, U_n^1 is contained in some V_δ^r, for $r = 1, 2, \cdots$, or h. For $U_n^1 \cdot V^r_{\delta/2} \neq 0$, for some r, since $n > M$. Also $\delta(U_n^1) < \delta/2$ since $n > N$. Hence $U_n^1 \subset V_\delta^r$. For any point $x \,\varepsilon\, (A - \sum_{i=N_\epsilon+1}^{\infty} U_i^1)$, $T_n(x) = T_{N_\epsilon} W_{N_\epsilon+1} \cdots W_n(x)$ and $T_{n+p}(x) = T_{N_\epsilon} W_{N_\epsilon+1} \cdots W_{n+p}(x)$, for $n > N_\epsilon$. Now all W_j, for $j > N_\epsilon$, are stationary outside of $\sum_{i=N_\epsilon+1}^{\infty} U_i^1$. Hence $T_n(x) = T_{N_\epsilon}(x) = T_{n+p}(x)$. Therefore, for any such x, $\rho(T_n(x), T_{n+p}(x)) = 0$, for $n > N_\epsilon$, and $p = 1, 2, \cdots$. For any $x \,\varepsilon \sum_{i=N_\epsilon+1}^{\infty} (U_i^1 - y_i)$, we have that $T_n(x) = T\, W_{N_\epsilon+1} \cdots W_n(x)$, since, by our definition of the neighborhoods U_i^1, all W_j, for $j \leq N_\epsilon$ are stationary for all such points x. Now x is contained in U_h^1, for some h; and U_h^1 is contained in V_δ^r, for some r. Each W_k, for $N_\epsilon < k \leq n$, transforms a point of U_h^1 only into some other point of U_h^1. Therefore $y = W_{N_\epsilon+1} \cdots W_n(x)$ is contained in U_h^1 which is contained in V_δ^r. Likewise

$$z = W_{N_\epsilon+1} \cdots W_{n+p}(x) \subset U_h^1 \subset V_\delta^r.$$

Hence

$$\rho(y, z) < \delta_{\epsilon/2}.$$

However

$$T_n(x) = T\, W_{N_\epsilon+1} \cdots W_n(x) = T(y)$$

and

$$T_{n+p}(x) = T\, W_{N_\epsilon+1} \cdots W_{n+p}(x) = T(z).$$

Accordingly

$$\rho(T_n(x), T_{n+p}(x)) = \rho(T(y), T(z)) < \epsilon, \quad \text{for } n > N_\epsilon, \text{ and } p = 1, 2, \cdots,$$

since T is uniformly continuous. For any y_i, $T_i(y_i) = b_i$, by definition. But $b_i = T(y_i)$. Hence $T_i(y_i) = T(y_i)$, for all i. If $x = y_i$, for any $i > N_\epsilon$, $T_n(x) = T_i W_{i+1} \cdots W_n(x)$, for $i > N_\epsilon$, $n > N_\epsilon$, and $n > i$. But all W_j, for $i < j \leq n$, are stationary on y_i. Therefore

$$T_n(x) = T_i(x) = T(x), \text{ for all } i < n.$$

If $n \leq i$, $T_n(x) = T(x)$, by definition. In the same way $T_{n+p}(x) = T(x)$, for $i > N_\epsilon$, $n > N_\epsilon$. Hence

$$\rho(T_n(x), T_{n+p}(x)) = 0, \text{ for } n > N_\epsilon,\ p = 1, 2, \cdots, \text{ for } x \subset \sum_{i=N_\epsilon+1}^{\infty} y_i.$$

Therefore, for any $\epsilon > 0$, there exists an N_ϵ such that for any $n > N_\epsilon$, and for any

$$x \,\varepsilon\, A,\ \rho(T_n(x), T_{n+p}(x)) < \epsilon, \text{ for } p = 1, 2, \cdots.$$

Hence $\{T_n\}$ converges uniformly to S. Consequently S is continuous.

Furthermore S is univalued both ways. For take any two distinct points p and q of A. There exists an integer N such that $p + q \subset A - \sum_{i=N}^{\infty} U_i^1$ Moreover T_N is univalued both ways over $A - \sum_{i=N}^{\infty} U_i^1$. Therefore

$$T_N(p) \neq T_N(q).$$

Now for any point

$$x \varepsilon A - \sum_{i=N}^{\infty} U_i^1, \; T_{N+j}(x) \equiv T_N(x), \text{ for all } j.$$

Hence

$$S(p) = T_{N+j}(p) = T_N(p) \neq T_N(q) = T_{N+j}(q) = S(q).$$

Accordingly $S(A) = B$ is a homeomorphism.

3°. We now assume that the lemma is true when $G_{k-1} = 0$ and proceed to prove that it is then true when $G_{k-1} \neq 0$ but $G_k = 0$.

Take a decomposition of A into the sets of G_{k-1} and the points of $A - G_{k-1}$. This is obviously an upper semi-continuous decomposition. Let its hyperspace be C and let $t(A) = C$ be the continuous transformation associated with the decomposition. The sets of G_{k-1} form a null sequence $\{g_i^k\}$ since $G_k = 0$. Let $t(g_i^k) = c_i \varepsilon C$, for each i. Now for any point $c \varepsilon C$, $t^{-1}(c)$ is a single point unless $c = c_i$, for some i; and $t^{-1}(c_i) = g_i^k = T^{-1}(b)$, a non-degenerate set, for some $b \varepsilon B$. Hence $\{t^{-1}(c_i)\}$ is the null sequence $\{g_i^k\}$. Furthermore, lim. sup. $\{t^{-1}(c_i)\} =$ lim. sup. $\{g_i^k\}$ which is L_{k-1}, and the collection of sets of G which intersect L_{k-1} is vacuous, by hypothesis. Moreover, for any $\epsilon > 0$, any $t^{-1}(c)$, for $c \varepsilon C$, and any $y \varepsilon t^{-1}(c)$, there exists a homeomorphism $W(A - y) = A - t^{-1}(c)$ which is stationary outside of the ϵ-neighborhood of $t^{-1}(c)$, because any non-degenerate $t^{-1}(c)$ is a set $T^{-1}(b)$, for some $b \varepsilon B$. Hence the conditions of this lemma are satisfied by $t(A) = C$ for the case demonstrated in part 2°. Therefore A is homeomorphic with C.

Let $Z(C) = T\,t^{-1}(C) = B$. Then $Z(C) = B$ is a univalued transformation. Moreover it is continuous, since T and t are each continuous, and A is compact. All $Z^{-1}(b)$, for $b \varepsilon B$, are degenerate except when $T^{-1}(b)$ is a set of $G - G_{k-1}$. Let H be the collection of all non-degenerate sets $Z^{-1}(b)$, for $b \varepsilon B$. Let H_1 be the collection of all sets of H which intersect $M =$ lim. sup. H. Let H_2 be the collection of all sets of H_1 which intersect $M_1 =$ lim. sup. H_1, etc. Wherefore, $H = t(G)$, and $H_i = t(G_i)$, for each $i < k$. However $t(G_{k-1})$ consists of the collection $\{c_i\}$ of single points of C. Hence $H_{k-1} = 0$.

Take any $\epsilon > 0$, any non-degenerate $Z^{-1}(b)$, for $b \varepsilon B$, and any point $y \varepsilon Z^{-1}(b)$. We must show that there exists a homeomorphism

$$W(C-y)=C-Z^{-1}(b)$$

which is stationary outside of the ϵ-neighborhood U_ϵ of $Z^{-1}(b)$ in C.

Now t is uniformly continuous, since A is compact. Hence there exists a δ_ϵ so that if $\rho(r,s)<\delta_\epsilon$, for $r,s\,\varepsilon\,A$, then $\rho(t(r),t(s))<\epsilon$. Take a neighborhood V of $T^{-1}(b)$ in A such that V is contained in the δ_ϵ-neighborhood of $T^{-1}(b)$, and such that $V\cdot g_i^k=0$, for all i. We can do this because the sets $T^{-1}(p)$, for $p\,\varepsilon\,B$, are closed and disjoint, and $T^{-1}(b)\cdot\bar{G}_{k-1}=0$, since $Z^{-1}(b)$ is non-degenerate in C. If we let $x=t^{-1}(y)$, then $x\,\varepsilon\,T^{-1}(b)$. Now, by hypothesis, there exists a homeomorphism $R(A-x)=A-T^{-1}(b)$ which is stationary in $A-V$. However, $A-x=t^{-1}(C-y)$. Hence

$$R\,t^{-1}(C-y)=A-T^{-1}(b).$$

We now designate by W the transformation

$$t\,R\,t^{-1}(C-y)=C-Z^{-1}(b).$$

Now $W(C-y)=C-Z^{-1}(b)$ is univalued. For take any point $q\,\varepsilon\,C-y$. If $q\neq c_i$, for any i, $W(q)=t\,R\,t^{-1}(q)$ is a single point of $C-Z^{-1}(b)$, because t is one-to-one for such points q and R is a homeomorphism. If $q=c_i$, for some i, then $t^{-1}(q)=g_i^k$. Furthermore $R\,t^{-1}(q)=t^{-1}(q)$, since $g_i^k\cdot V=0$ and R is stationary outside of V. Therefore $W(q)=t\,t^{-1}(q)=q$. By a similar proof we see that W^{-1} is also univalued. Hence W is one-to-one.

Furthermore, W and W^{-1} are both continuous since t and R are continuous and A is compact. Therefore W is a homeomorphism.

Now W is stationary outside of U_ϵ. For if we take any point $q\,\varepsilon\,C-U_\epsilon$, then $t^{-1}(q)\cdot V=0$. Hence, since R is stationary outside of V,

$$R\,t^{-1}(q)=t^{-1}(q).$$

Thus we have

$$W(q)=t\,R\,t^{-1}(q)=t\,t^{-1}(q)=q.$$

Accordingly the homeomorphism

$$W(C-y)=C-Z^{-1}(b)$$

is stationary outside of U_ϵ.

Hence the conditions of the lemma are satisfied by $Z(C)=B$ for the case $k-1$. We have assumed that for this case the lemma is true. Wherefore, C is homeomorphic with B. However, we have seen that A is homeomorphic with C. Therefore A is homeomorphic with B.

6. *Corollaries.* There are some properties of the homeomorphism, say $S(A) = B$, whose existence is established in the above lemma, which follow from the proof of the lemma. These will be stated and verified in the following two corollaries. We will assume that the conditions of the lemma are satisfied.

COROLLARY I. *The homeomorphism $S(A) = B$, whose existence is established in the lemma, is such that $S \equiv T$ over all points of A lying in $A - \sum_{j=1}^{k} \sum_{i=1}^{\infty} U_i^j$.*

Proof. We shall make the proof by induction. Take any point p in $A - \sum_{j=1}^{k} \sum_{i=1}^{\infty} U_i^j$. If $k = 1$, then $\{g_i^1\}$ is a null sequence. Now $W_j(p) = p$, for all j, since W_j is stationary outside of U_j^1, for every j. Now, for all m, $T_m(p) = T\, W_1 \cdots W_m(p)$. Hence $T_m(p) = T(p)$, for all m. Therefore $S(p) \equiv T(p)$ for this case.

Now let us assume that this is also true for $k - 1$, and prove that it is then true for k. The sequence $\{g_i^k\}$ is a null sequence, since $G_k = 0$. Take a decomposition of A into the sets of $\{g_i^k\}$ and the points of $A - \sum_{i=1}^{\infty} g_i^k$. This is an upper semi-continuous decomposition. Let its hyper-space be C and let $t(A) = C$ be the continuous transformation associated with the decomposition. Now $t(A) = C$ satisfies the conditions of the lemma for the case $k = 1$. Hence there exists a homeomorphism $R_1(A) = C$ so that $R_1 \equiv t$ over all points of the set $A - \sum_{i=1}^{\infty} U_i^k$, as we have just demonstrated above. Now we established in part 3° of the proof of the lemma that the conditions of the lemma are satisfied for the case $k - 1$ by the spaces C and B and the transformation $Z(C) = T\, t^{-1}(C) = B$, where the non-degenerate sets in C are the sets $t(g_i^j)$ and the corresponding neighborhoods are the neighborhoods $t(U_i^j)$ for all i, and for $j < k$. From our assumption it then follows that there exists a homeomorphism $R_2(C) = B$ such that $R_2 \equiv Z$ over $C - \sum_{j=1}^{k-1} \sum_{i=1}^{\infty} t(U_i^j)$. Now $S(A) = R_2 R_1(A) = B$. Furthermore $S(p) = R_2 R_1(p) = R_2 t(p)$, since $R_1(p) \equiv t(p)$. Now $t(p) \subset C - \sum_{j=1}^{k-1} \sum_{i=1}^{\infty} t(U_i^j)$. Hence $R_2 t(p) = Z\, t(p)$. However $Z\, t(p) = T\, t^{-1} t(p) = T(p)$. Thus $S(p) = T(p)$ for any point p in $A - \sum_{j=1}^{k} \sum_{i=1}^{\infty} U_i^j$.

COROLLARY II. *The homeomorphism $S(A) = B$, whose existence is established in the lemma, is such that, for any $p \in A$, if $p \subset U_d^a$, for any*

$\alpha \leq k$, and any d, $S^{-1}T(p) \subset U_r^\rho$, for some $\rho \leq k$ and some r, such that there is a chain of neighborhoods U_d^α, $U_e^\beta, \cdots, U_r^\rho$ so that each two consecutive neighborhoods in this chain intersect each other.

Proof. We shall make the proof by induction.

If $k = 1$, then $\alpha = 1$. Now, for any j, $S(U_j^1) = T_j(U_j^1)$, since, for any point $x \varepsilon U_j^1$, $T_n(x) = T_j(x)$ for all $n > j$. Furthermore

$$T_j(U_j^1 - y_j) = T\,W_j(U_j^1 - y_j)$$

by definition, and because all W_i are stationary over U_j^1, for $i < j$, since $U_j^1 \cdot U_i^1 = 0$ for $i \neq j$. Now, since W_j is stationary in $A - U_j^1$, we have

$$W_j(U_j^1 - y_j) = U_j^1 - g_j^1.$$

Hence we have

$$S(U_j^1) - S(y_j) = S(U_j^1 - y_j) = T(U_j^1 - g_j^1) = T(U_j^1) - T(g_j^1).$$

But $S(y_j) = b_j = T(g_j^1)$. Therefore $S(U_j^1) = T(U_j^1)$. Thus, for any point $p \subset U_j^1$, for any j, $S^{-1}T(p) \subset U_j^1$.

We now assume that the statement is true for $k - 1$ and proceed to prove it for k. The sequence $G_{k-1} = \{g_i^k\}$ is a null sequence, since $G_k = 0$. Take a decomposition of A into the sets of G_{k-1} and the points of $A - G_{k-1}$, and thus obtain the hyperspace C, the transformations

$$t(A) = C,\ Z(C) = T\,t^{-1}(C) = B$$

and the homeomorphisms

$$R_1(A) = C,\ R_2(C) = B \text{ and } S(A) = R_2R_1(A) = B$$

as in the proof of Corollary I. From our assumption it follows that $R_2(C) = B$ is such that $R_2^{-1}Z(C) = C$ satisfies the results of this corollary for the case $k - 1$. Now $T(A) = Z\,t(A) = B$. Hence we have that

$$S^{-1}T(A) = R_1^{-1}R_2^{-1}Z\,t(A) = A.$$

Take any point p of A. If $p \subset U_i^j$, for any $j \leq k$, and any i, then

$$q = t(p) \subset t(U_i^j).$$

According to our assumption, if $q \subset t(U_d^\alpha)$, for any $\alpha \leq k - 1$, and any d, $R_2^{-1}Z(q)$ is in a neighborhood $t(U_r^\rho)$ for some ρ and some r, such that there is a chain of neighborhoods $t(U_d^\alpha), t(U_e^\beta), \cdots, t(U_r^\rho)$ such that each two consecutive neighborhoods in this chain intersect each other. Now if

$$R_2^{-1}Z(q) \not\subset \sum_{i=1}^{\infty} t(U_i^k),$$

then

$$S^{-1}T(p) = R_1^{-1}[R_2^{-1}Z(q)] \subset U_r,$$

since

$$R_1(x) = t(x) \text{ for any } x \in A - \sum_{i=1}^{\infty} U_i^k.$$

Now the chain of neighborhoods $U_i^j, U_d^a, \cdots, U_r^\rho$ is such that each two consecutive neighborhoods of the chain intersect each other. For $U_i^j \cdot U_d^a \neq 0$ since $t(U_i^j) \cdot t(U_d^a) \supset q$; and any other two consecutive ones intersect because their images under t do. Hence the corollary is established for this case. Now if $R_2^{-1}Z(q) \subset t(U_m^k)$, for some m, then we have that

$$S^{-1}T(p) = R_1^{-1}[R_2^{-1}Z(q)] \subset U_m^k,$$

since, for all m, $R_1(U_m^k) = t(U_m^k)$, by the first part of this proof. The chain $U_i^j, U_d^a, \cdots, U_r^\rho, U_m^k$ is obviously such that each two consecutive neighborhoods of the chain intersect. For

$$U_r^\rho \cdot U_m^k \neq 0 \text{ since } t(U_r^\rho) \cdot t(U_m^k) \supset R_2^{-1}Z(q),$$

and by the above argument. Hence the corollary is also established for this case.

7. THEOREM. *If A and B are compact metric spaces and $T(A) = B$ is continuous and satisfies the conditions:* (1) *for any $\epsilon > 0$, any set $T^{-1}(b)$, for $b \in B$, and any $x \in T^{-1}(b)$, there exists a homeomorphism*

$$W(A - x) = A - T^{-1}(b)$$

which is stationary outside of the ϵ-neighborhood of $T^{-1}(b)$; (2) *there exists some number α of the first or second number class such that $G_\alpha = 0$; and* (3) $\prod_0^\infty L_i$ *is a zero-dimensional set, then A is homeomorphic with B.*

Proof. The theorem is true for any finite number α, by the lemma. We will now prove that the theorem is true when $\alpha = \omega$. When this is established, it will follow that it is true for any $\alpha > \omega$. For, if α is an isolated number, the result is obtained by induction as in the lemma. If α is a limit number, we again use induction, that is, we suppose the theorem true for all $\beta < \alpha$, and then prove that it is also true for α by the same method of proof we shall use here for $\alpha = \omega$.

Let H_1 be the collection of all sets g_j^1 of $G - G_1$ such that $U_j^1 \cdot \sum_{r=2}^{\infty} \sum_{i=1}^{\infty} U_i^r = 0$. Let $\{h_i^1\}$ represent the collection H_1. Let V_i^1 denote the neighborhood $\bar{U}_k^1$ which covers h_i^1, for each i. Let H_2 be the collection of all sets of $(G - H_2) - G_2$ whose neighborhoods U do not intersect $\sum_{r=3}^{\infty} \sum_{i=1}^{\infty} U_i^r$. Let $\{h_i^2\}$ represent the collection H_2; and let V_i^2 denote the neighborhood U_k^s, $s = 1$ or 2, which covers h_i^2, for each i. Let H_3 be the collection of all sets of $(G - \sum_{k=1}^{2} H_k) - G_3$ whose neighborhoods U do not intersect $\sum_{r=4}^{\infty} \sum_{i=1}^{\infty} U_i^r$. Let $\{h_i^3\}$ denote the collection H_3; and let V_i^3 denote the neighborhood U_k^s. $s = 1, 2$, or 3, which covers h_i^3, for each i. Continue in this way indefinitely.

Now $\sum_{i=1}^{\infty} H_i = G$. For take any $g \varepsilon G$. There exists a finite integer N which is the largest number so that $g \varepsilon G_N$. Now, by definition, the neighborhood U of g is contained in the $\rho(g, L_N)/2$-neighborhood of g, thus $\bar{U} \cdot L_N = 0$. Take a neighborhood V of L_N so that $V \cdot \bar{U} = 0$. Now there exists an integer M so that $U_i^m \subset V$, for all $m > M$ and all i, since lim. sup. $\sum_{j=N+1}^{\infty} \sum_{i=1}^{\infty} U_i^j = L_N$. Therefore U does not intersect $\sum_{r=M+1}^{\infty} \sum_{i=1}^{\infty} U_i^r$. Hence $g \varepsilon H_s$ for some $s \leq M$.

Take a decomposition of A into the sets of H_1 and the points of $A - H_1$. This is obviously an upper semi-continuous decomposition. Let its hyperspace be C_1 and let $T_1(A) = C_1$ be the continuous transformation associated with the decomposition. Now no sets of H_1 intersect lim. sup. H_1. Therefore $T_1(A) = C_1$ satisfies the conditions of the lemma for the case $k = 1$. Hence, by Corollary I section 6, there exists a homeomorphism $S_1(A) = C_1$ such that $S_1 \equiv T_1$ over $A - \sum_{i=1}^{\infty} V_i^1$. Let $Z_1(A) = S_1^{-1}T_1(A) = A$. Now Z_1 is univalued and continuous; $Z_1(h_i^1)$ is a point of A, for each i; and $Z_1(h_i^k) = h_i^k$, for all $k > 1$ and for all i.

Take a decomposition of A into the sets of H_2 and the points of $A - H_2$. This is an upper semi-continuous decomposition. Let C_2 be its hyperspace and $T_2(A) = C_2$ its associated continuous transformation. Now $T_2(A) = C_2$ satisfies the conditions of the lemma for the case $k = 2$. Hence there exists a homeomorphism $S_2(A) = C_2$ such that $S_2 \equiv T_2$ over $A - \sum_{i=1}^{\infty} V_i^2$. Let $Z_2(A) = S_2^{-1}T_2Z_1(A) = A$. Now Z_2 is univalued and continuous and such that: $Z_2(h_i^j)$ is a point of A, for $j = 1$ or 2 and for each i; and Z_2 is the identity transformation over $A - \sum_{k=1}^{2} \sum_{i=1}^{\infty} V_i^k$. Hence $Z_2(h_i^k) = h_i^k$, for all

$k > 2$ and for all i. Continuing in this way indefinitely we get a sequence $\{Z_n\}$ of such transformations.

Let $Z = \lim_{n=\infty} \{Z_n\}$. We shall now prove that $\{Z_n\}$ converges uniformly to Z. It will then follow that Z is continuous. Take any $\epsilon > 0$. Now since $\prod_0^\infty L_i$ is closed, compact, and zero-dimensional, we can cover it by a finite number of disjoint $\epsilon/2$-neighborhoods, say $W_1, W_2, \cdots, W_h$. Let W denote $\sum_{i=1}^h W_i$. Now there exists an M such that $\sum_{j=M+1}^\infty \sum_{i=1}^\infty V_i{}^j \subset W$. For, if not, then no matter what M we take, there exists some $m > M$ and some i so that $V_i{}^m \not\subset W$. Take a sequence $M_1 < M_2 < M_3 < \cdots$ of integers and take the corresponding $V_i{}^{m_j} \not\subset W$. Take a point p_{m_j} of $V_i{}^{m_j}$ but not contained in W. Now $\{p_{m_j}\}$ contains a convergent subsequence $\{p'_{m_j}\}$ which converges to a point p. Now $p \subset L$ because lim. sup. $\sum_{j=1}^\infty \sum_{i=1}^\infty V_i{}^j = L$. Also, for any finite integer n, infinitely many of the points of $\{p'_{m_j}\}$ have subscripts greater than n. Therefore $p \subset L_n$ for every finite n. Hence $p \subset \prod_0^\infty L_i$, which is a contradiction. Furthermore there exists an integer $N \geq M$ so that for all $n > N$ and all i, $V_i{}^n \subset W_r$ for some r. For, if not, then no matter what N we take, there exists some $n > N$ and some i so the $V_i{}^n \not\subset W_r$, for any r. Take a sequence $\{N_j\}$ of integers such that $M < N_1 < N_2 < N_3 < \cdots$, and take the corresponding $V_{i_j}^{n_j} \not\subset W_r$, for any r. Now no $V_{i_j}^{n_j}$ will intersect both W_s and W_r, for $r, s = 1, 2, \cdots$, or h, unless the corresponding set $h_{i_j}^{n_j}$ intersects both W_s and W_r, since $W_i \cdot W_j = 0$, for $i \neq j$, and since each $V_{i_j}^{n_j} \subset W$. Hence there exists some u and some v so that each set of some subcollection $\{h_k\}$ of the collection $\{h_{i_j}^{n_j}\}$ have points in both W_u and W_v. Take a point p_k in $h_k \cdot W_u$ and a point q_k in $h_k \cdot W_v$, for each k. Now $\{p_k\}$ contains a convergent subsequence $\{p'_k\}$ which converges to a point $p \subset \prod_0^\infty L_i$, and $\{q_k\}$ contains a convergent subsequence $\{q'_k\}$ which converges to a point $q \subset \prod_0^\infty L_i$. Obviously p and q are distinct points. Furthermore, since p'_k and q'_k are on the same set $h_k = T^{-1}(b_k)$, for some $b_k \varepsilon B$, and since T is continuous, it follows that $p + q \subset T^{-1}(b)$, for some $b \varepsilon B$. Hence this set $T^{-1}(b)$ is a set of G_ω, which contradicts our assumption that $G_\omega = 0$.

The integer N depends upon ϵ. Take any $x \varepsilon A$. Now

$$Z_n(x) = S_n^{-1} T_n \cdots S_{N+1}^{-1} T_{N+1} Z_N(x)$$

and

$$Z_{n+p}(x) = S_{n+p}^{-1} T_{n+p} \cdots S_n^{-1} T_n \cdots S_{N+1}^{-1} T_{N+1} Z_N(x).$$

If $Z_N(x) \not\subset \sum_{j=N+1}^{\infty} \sum_{i=1}^{\infty} V_i^j$, then $S_k^{-1} T_k Z_N(x) = Z_N(x)$, for all $k > N$. Therefore $Z_n(x) = Z_N(x)$. Likewise $Z_{n+p}(x) = Z_N(x)$, for $p = 1, 2, \cdots$. Hence $\rho(Z_n(x), Z_{n+p}(x)) = 0$, for $n > N$ and $p = 1, 2, \cdots$, in this case. If $Z_N(x) \subset \sum_{j=N+1}^{\infty} \sum_{i=1}^{\infty} V_i^j$, then $Z_N(x) \subset W_r$ for some $r = 1, 2, \cdots$, or h. If $Z_N(x) \subset V_{d_1}^{N+1}$, for some d_1, then $S_{N+1}^{-1} T_{N+1} Z_N(x)$ is a point of $V_{m_1}^{N+1}$, for some m_1, where $V_{m_1}^{N+1} \subset W_r$. For, by Corollary II, there is a chain of neighborhoods $V_{d_1}^{N+1}, V_{e_1}^{N+1}, \cdots, V_{m_1}^{N+1}$, so that each two consecutive neighborhoods in the chain intersect. Hence $V_{m_1}^{N+1} \subset W_r$ since none of the V^{N+1}'s lying in any of the other W_i's will intersect those in W_r, as $W_i \cdot W_j = 0$, for $i \neq j$. Therefore $S_{N+1}^{-1} T_{N+1} Z_N(x) \subset W_r$ for this case. If $Z_N(x) \not\subset \sum_{i=1}^{\infty} V_i^{N+1}$, then we have that $S_{N+1}^{-1} T_{N+1} Z_N(x) = Z_N(x) \subset W_r$. If $S_{N+1}^{-1} T_{N+1} Z_N(x) \subset V_{d_2}^{N+2}$, for some d_2, then $S_{N+2}^{-1} T_{N+2} S_{N+1}^{-1} T_{N+1} Z_N(x)$ is a point of $V_{m_2}^{N+2}$, for some m_2, where $V_{m_2}^{N+2} \subset W_r$, by the same argument as above. If $S_{N+1}^{-1} T_{N+1} Z_N(x) \not\subset \sum_{i=1}^{\infty} V_i^{N+2}$, then we have that $S_{N+2}^{-1} T_{N+2} S_{N+1}^{-1} T_{N+1} Z_N(x) = S_{N+1}^{-1} T_{N+1} Z_N(x) \subset W_r$. Continue in this same way. Finally if the set $S_{n-1}^{-1} T_{n-1} \cdots S_{N+1}^{-1} T_{N+1} Z_N(x) \subset V^n_{d_n}$, for some d_n, then $S_n^{-1} T_n S_{n-1}^{-1} T_{n-1} \cdots Z_N(x) \subset V^n_{m_n}$, for some m_n, where $V^n_{m_n} \subset W_r$. If

$$S_{n-1}^{-1} T_{n-1} \cdots S_{N+1}^{-1} T_{N+1} Z_N(x) \not\subset \sum_{i=1}^{\infty} V_i^n$$

then

$$S_n^{-1} T_n S_{n-1}^{-1} T_{n-1} \cdots Z_N(x) = S_{n-1}^{-1} T_{n-1} \cdots Z_N(x) \subset W_r.$$

Therefore $Z_n(x)$ is a point p of W_r, and likewise

$$Z_{n+p}(x) = S_{n+p}^{-1} T_{n+p} \cdots S_{n+1}^{-1} T_{n+1}(p) = q \in W_r.$$

Accordingly

$$\rho(Z_n(x), Z_{n+p}(x)) = \rho(p, q) < \epsilon, \text{ for } n > N, p = 1, 2, \cdots;$$

and so $\{Z_n\}$ converges uniformly to Z. Hence Z is continuous.

Now Z transforms any set $g \in G$ into a single point of A. Furthermore, for any point $a \in A$, $Z^{-1}(a)$ is either a point p of A or a set $g \in G$ in A. Consider $T Z^{-1}(A) = B$. If $Z^{-1}(a) = p \in A$, $T Z^{-1}(a) = T(p) \in B$. If $Z^{-1}(a) = g$, where $g = T^{-1}(b)$, for some $b \in B$, then $T Z^{-1}(a) = T(g)$ which

is the point b. Hence $T Z^{-1}(A) = B$ is univalued. It is obviously continuous, since T and Z are continuous and A is compact.

Now $[T Z^{-1}]^{-1}(B) = Z T^{-1}(B) = A$ is also univalued. For take any point $b \varepsilon B$. If $T^{-1}(b) = p \varepsilon A$, then $Z T^{-1}(b) = Z(p)$ which is a point of A. since Z is univalued. If $T^{-1}(b) = g \varepsilon G$ in A, then $Z T^{-1}(b) = Z(g)$ which is also a point of A by the definition of Z.

Therefore $T Z^{-1}(A) = B$ is continuous and is univalued both ways. Hence, since A is compact, it is a homeomorphism.

8. *Applications.* The following two theorems classify some of the types of sets into which we can decompose the compact Euclidean 3-space,[12] so that, if the decomposition is upper semi-continuous and conditions (2) and (3) of the last theorem are satisfied, the hyperspace of the decomposition is also a compact Euclidean 3-space.

THEOREM a. *Let A be a compact Euclidean 3-space and $T(A) = B$ a continuous transformation such that each non-degenerate set $T^{-1}(b)$, for $b \varepsilon B$, is an isotopic sphere[13] plus its interior. Define the collections G and G_i and the sets L_i, for each i, as before. If there exists a number α of the first or second number class so that $G_\alpha = 0$; and if $\prod_0^\infty L_i$ is a zero-dimensional set; then B is also a compact Euclidean 3-space.*

Proof. Take any $\epsilon > 0$, any set $g \varepsilon G$, any $x \varepsilon g$, and any geometric sphere C in A. Let $S = C +$ its interior. There exists a homeomorphism $Z(A) = A$ so that $Z(S) = g$, since g is an isotopic sphere plus its interior. Let $Z^{-1}(x) = y$. Now $y \subset S$. It is well known that there exists a homeomorphism $R(A - y) = A - S$ which is stationary outside of the δ_ϵ-neighborhood U_δ of S. Let $W(A - x) = Z R Z^{-1}(A - x) = A - g$. Now W is a homeomorphism which is stationary outside of $V_\epsilon(g)$ by the proof used in the lemma for similar transformations. Therefore the conditions of the theorem in the last section are satisfied. Accordingly A is homeomorphic with B.

THEOREM b. *Let A be a compact Euclidean 3-space and $T(A) = B$ a continuous transformation such that each non-degenerate $T^{-1}(b)$, for $b \varepsilon B$, is a convex set in A. Define the collections G and G_i, and the sets L_i, for each i, as before. If there exists a number α of the first or second number*

[12] That is, the ordinary Euclidean 3-space compactified by adding the points at infinity.

[13] A set S in A is said to be an *isotopic sphere* if, for any geometric sphere C in A there exists a homeomorphism $R(A) = A$ such that $R(C) = S$.

class so that $G_d = 0$; and if $\prod_0^\infty L_i$ is a zero-dimensional set: then B is also a compact Euclidean 3-space.

Proof. We again take any $\epsilon > 0$, any set $g \varepsilon G$, and any $x \varepsilon g$ in A. Also take a convex ϵ-neighborhood V_ϵ of g in A. Take a system of spherical coördinates in A with the point x as origin. Choose any θ and any ϕ. The ray determined by this choice will intersect $F(V_\epsilon)$ in a single point, say (d, θ, ϕ), and it will intersect $F(g)$ in a single point, say (c, θ, ϕ). On this ray set up the following homeomorphism: $r' = \dfrac{(d-r)\cdot c}{d} + r$, if $r \leq d$, and $r' = r$, if $r > d$, where (r, θ, ϕ) represents any point on this ray. If we let θ and ϕ vary over all possible values, we get a homeomorphism $W(A - x) = A - g$ which is stationary outside of V_ϵ. Therefore, by the theorem of section 7, A is homeomorphic with B.

Since conditions I and III of section 2 were shown in section 3 to be equivalent when A is a 2-dimensional manifold, and in view of the theorem of section 7, we have the following:

THEOREM c. *The hyperspace B of any upper semi-continuous decomposition G of a 2-dimensional manifold A, where G is such that, for any $g \varepsilon G$ and any $x \varepsilon A$, there exists a homeomorphism $W(A - x) = A - g$, and such that conditions* (2) *and* (3) *of the theorem of section 7 are satisfied, is the same kind of 2-dimensional manifold, i. e. is homeomorphic with A.*

THE JOHNS HOPKINS UNIVERSITY.

ERGODIC CURVES.

By Monroe H. Martin.

The problem of determining the least time required by the point of some motion to come within a distance ϵ of every point of the phase space was first considered by Birkhoff.[1] More recently Errera[2] has treated the problem in the plane of determining the length of the shortest curve, such that the distance of every point of a given domain D is at a distance $\leqq \epsilon$ from some point of the curve. He has solved the problem in case the domain D can be "swept out" by moving the center of a circle of radius ϵ along a Jordan curve of a certain special type. However, his methods fail to yield any information for an arbitrary domain D and it is not even certain, *a priori*, that a curve of shortest length, possessing the required property, exists.

In this paper this existence question is taken up in a slightly more general form, the domain D being replaced by a bounded point set M. The set of continuous, rectifiable curves, which are such that an arbitrary point of M lies at a distance $\leqq \epsilon$ from some point of the curve, is proved to contain at least one curve whose length equals the greatest lower bound, $\Lambda(\epsilon)$, of the lengths of the curves in the set. The set of such curves is shown to be closed. $\Lambda(\epsilon)$ is proved to be a monotone non-increasing function of ϵ which is continuous on the right for any positive ϵ. Whether $\Lambda(\epsilon)$ is continuous on the left, is an open question. The relation of the behavior of $\Lambda(\epsilon)$ in the neighborhood of $\epsilon = 0$ to the structure of M is investigated.

The author is indebted to Professor Tamarkin for pointing out the usefulness of a result of Tonelli (Lemma 4 in the present paper), the author's own proof having been invalid at this point.

1. *Notation and definitions.* Unless explicitly stated otherwise, ϵ denotes a positive number throughout the paper. x, y denote ordinary Cartesian coördinates in the Euclidean plane E, and M a bounded point set in E containing more than one point of E. The Euclidean distance between two points P, Q of E is denoted by $|PQ|$.

If P is a fixed point in E, the set of those points Q in E for which

[1] G. D. Birkhoff, "Probability and physical systems," *Bulletin of the American Mathematical Society*, vol. 38 (1932), pp. 375-377.

[2] A. Errera, "Un problème de géométrie infinitésimale," *Memoires Académie royale de Belgique*, vol. 12, no. 1441 (1932).

$|PQ| \leqq \rho$ will constitute a *circular region.* ρ is the *radius* of the circular region and P its *center*. A circular region Γ *encloses* M if every point of M is a point of Γ.

A *continuous curve*[3] is an ordered set of points in E defined by the equations

$$x = f(t), \qquad y = g(t),$$

for t varying from a to b, where $f(t)$, $g(t)$ are real, continuous, single-valued functions of t defined in the closed interval $[a, b]$. If $f(t)$, $g(t)$ have bounded variation in $[a, b]$, the curve is *rectifiable.* A continuous curve is *ϵ-ergodic to* M if an arbitrary point of M is at a distance $\leqq \epsilon$ from some point of the curve. The set of points on a continuous curve is bounded, perfect and connected. In particular, it contains all of its limit points. Hence, if a continuous curve C be ϵ_m-ergodic to M and $\lim\limits_{m\to\infty} \epsilon_m = 0$, it follows that C contains all the points of M. More generally, if C be ϵ_m-ergodic to M and $\lim\limits_{m\to\infty} \epsilon_m = \epsilon$, it is seen that C is ϵ-ergodic to M. Two continuous curves

$$C_i: \qquad x = f_i(t), \qquad y = g_i(t), \qquad a \leqq t \leqq b, \qquad (i = 1, 2),$$

will be said to *lie in the η-neighborhood of each other,*[4] if the points on them corresponding to the same value of t always lie at a distance $\leqq \eta$ from each other.

Consider a succession

$$A_1, A_2, \cdots, A_n, \cdots,$$

of sets of points in E. A point P of E will be a *point of accumulation of the succession*[5] if for any positive number η, there are infinitely many sets A_n of the succession having at least one point P_n satisfying the condition $|PP_n| \leqq \eta$. Given a set of continuous curves

$$S: \qquad \{x = f(t), y = g(t)\}, \qquad 0 \leqq t \leqq 1,$$

a continuous curve

$$C: \qquad x = F(t), \qquad y = G(t), \qquad 0 \leqq t \leqq 1,$$

is a *curve of accumulation of S*,[6] if for any positive number η, there are infinitely many curves C' of S such that C, C' lie in the η-neighborhood of each other. S is said to be *closed* if it contains all of its curves of accumulation.

---

[3] Cf., for example, L. Tonelli, *Fondamenti di calcolo delle variazioni*, vol. 1, p. 34.

[4] Cf. Tonelli, *op. cit.*, p. 72.

[5] Cf. M. Cipolla, "Sul postulato di Zermelo e la teoria dei limiti delle funzioni," *Atti dell'Accademia Gioenia in Catania*, Serie 5ª, vol. 6, p. 3, or Tonelli, *op. cit.*, p. 73.

[6] Tonelli, *op. cit.*, pp. 73-74.

Given a succession

$$S_1, S_2, \cdots, S_n, \cdots,$$

of sets of continuous curves, all defined in the interval $0 \leqq t \leqq 1$, the curve C is a *curve of accumulation of the succession*,[6] if for any positive number η, there are infinitely many sets S_n of the succession having at least one curve C_n such that C, C_n lie in the η-neighborhood of each other. The curve C will be a *limit curve of the succession*,[7] if for any positive number η, there exists an N_η such that C and an arbitrary member of S_n lie in the η-neighborhood of each other for $n > N_\eta$. The succession is then said to *converge uniformly towards* C.[7]

The adoption of the above definitions enables us later on to avoid the use of Zermelo's general selection principle.

2. *Preliminary lemmas.* In this section we collect four lemmas for later reference.

LEMMA 1. *Among the circular regions enclosing a given M, there is one, Γ, whose radius ρ $(\rho > 0)$ is less than that of any other.*

Let $F(P)$ denote the least upper bound of $|PQ|$, where P is a fixed point in E and Q an arbitrary point of M. $F(P)$ is defined and continuous at any P. The value of the greatest lower bound of $F(P)$ is taken at least once by $F(P)$, say at R.

A circular region of radius $F(R)$ with R as center encloses M, and is the circular region Γ of least radius. For, $F(P) \geqq F(R)$, and the equality sign holds only if P coincides with R. To prove this, suppose $F(P) = F(R)$ with P and R different. Take P and R as centers of circular regions of radii $F(R)$. The two circular regions so obtained must intersect, since each encloses M. The part common to the two circular regions contains M and can be enclosed by a circular region of radius less than $F(R)$. But this is impossible, since $F(R)$ is the greatest lower bound of $F(P)$.

LEMMA 2. *If a succession $S'_1, S'_2, \cdots, S'_m, \cdots$ of sets of continuous curves ϵ-ergodic to M converges uniformly towards a limit curve C, then C is also ϵ-ergodic to M.*[8]

---

[7] Tonelli, *op. cit.*, p. 76.

[8] When the convergence is not uniform, C is not necessarily ϵ-ergodic to M. For example, let M comprise the points of the two intervals on $y = 0$, $y = 1$ for which $0 \leqq x \leqq \epsilon$ $(\epsilon < 1/2)$, and let S'_m contain the single curve

$$x = t, \qquad y = em^2\epsilon^{-2}t^2e^{-m^2\epsilon^{-2}t^2}, \qquad 0 \leqq t \leqq 1.$$

Here C is

$$x = t, \qquad y = 0, \qquad 0 \leqq t \leqq 1,$$

and is not ϵ-ergodic to M.

On the other hand C may be ϵ-ergodic to M even if the convergence be not uniform, as is the case in the example just given, when the points on $y = 1$ are removed from M.

Let P be an arbitrary point of M and γ be the circular region with center P and radius ϵ. Denote by Σ_m the set of points in which γ is intersected by the members of S'_m. Since no Σ_m is vacuous, the succession $\Sigma_1, \Sigma_2, \cdots, \Sigma_m, \cdots$ possesses [9] at least one point of accumulation on γ, say Q. If η denotes an arbitrarily small positive number, there are infinitely many sets Σ_m each having at least one point P_m, such that $|QP_m| < \eta/2$. If m be taken sufficiently large, due to the hypothesis of uniform convergence, there is a point R_m on C such that $|P_mR_m| < \eta/2$. Hence $|QR_m| < \eta$ and Q is a limit point for the points comprising C. Q therefore belongs to C, which establishes the lemma.

Lemmas 3 and 4 are due to Tonelli.[10]

LEMMA 3. *In a set S of infinitely many continuous, rectifiable curves [succession $S_1, S_2, \cdots, S_n, \cdots$ of sets of continuous, rectifiable curves], all contained in a finite domain and all having a length less than a fixed number, the following hold:*

(a) *there exists a parametric representation of the type*

$$x = f(t), \qquad y = g(t), \qquad 0 \leqq t \leqq 1,$$

for each curve of the set [succession];

(b) *the set [succession] possesses at least one continuous and rectifiable curve of accumulation*

$$C: \qquad x = F(t), \qquad y = G(t), \qquad 0 \leqq t \leqq 1;$$

(c) *there exists a process which makes a well-defined curve of accumulation C correspond to the given set [succession] and which determines a succession $S'_1, S'_2, \cdots, S'_m, \cdots$ which converges uniformly towards C, where S'_m is a set of curves extracted from $S[S_n(n \geqq m)]$.*

LEMMA 4. *For a set of continuous, rectifiable curves all of whose lengths are less than L, the curves of accumulation of the set are all rectifiable and have lengths not greater than L.*

3. *Ergodic curves and the ergodic function* $\Lambda(\epsilon)$. We begin with

DEFINITION 1. The greatest lower bound of the lengths of the con-

---

[9] M. Cipolla, *loc. cit.*, p. 3, proves this for linear point sets. The proof for point sets in the plane is quite simple and is omitted.

[10] *Op. cit.* For Lemma 3 see pp. 86-92, for Lemma 4 see p. 75.

tinuous, rectifiable curves ϵ-ergodic to M is the *ergodic function of M* and is denoted by $\Lambda(\epsilon)$.

We now prove

THEOREM 1. *The ergodic function $\Lambda(\epsilon)$ of M is finite and non-negative, being equal to zero if, and only if $\epsilon \geqq \rho$, where ρ is the radius of the circular region Γ in Lemma 1. There exists at least one well-defined continuous, rectifiable curve of length $\Lambda(\epsilon)$ which is ϵ-ergodic to M.*

To prove that $\Lambda(\epsilon)$ is finite, take r, θ as polar coördinates with the pole at the center of the circular region Γ in Lemma 1 and draw the curve $r = \epsilon\theta/2\pi$ $(0 \leqq \theta \leqq 2\pi\rho/\epsilon)$. This curve is ϵ-*ergodic* to M, and since it has finite length, $\Lambda(\epsilon)$ is finite.

Obviously $\Lambda(\epsilon)$ cannot be negative. That $\Lambda(\epsilon) = 0$ for $\epsilon \geqq \rho$ is trivial in view of the significance ascribed to ρ in Lemma 1. The proof that $\Lambda(\epsilon) > 0$ for $\epsilon < \rho$ is deferred until the end of the proof of this theorem.

Suppose $\epsilon < \rho$. We show that there exists a well-defined, continuous, rectifiable curve, ϵ-ergodic to M, whose length equals $\Lambda(\epsilon)$. Let $\lambda_1, \lambda_2, \cdots, \lambda_n, \cdots$ be a monotone decreasing sequence of positive numbers converging to $\Lambda(\epsilon)$. Consider the succession

$$S_1, S_2, \cdots, S_n, \cdots, \tag{1}$$

of sets of continuous, rectifiable curves ϵ-ergodic to M, in which S_n comprises the totality of such curves whose lengths are less than λ_n, but not less than λ_{n+1}. A circular region concentric with the circular region Γ in Lemma 1 and having a radius $\rho + \lambda_1 + \epsilon$ contains all the curves in (1), inasmuch as the lengths of the curves in (1) are all less than λ_1. Hence (1) can be identified as the succession $S_1, S_2, \cdots, S_n, \cdots$ in Lemma 3. According to this lemma, the succession possesses at least one well-defined, continuous, rectifiable curve of accumulation C, part (c) of the lemma providing a succession

$$S'_1, S'_2, \cdots, S'_m, \cdots, \qquad S'_m \subset S_n \ (n \geqq m),$$

of sets of curves which converges uniformly to C. Since the curves in the succession are all ϵ-ergodic to M, Lemma 2 applies to show that C is ϵ-ergodic to M. Now consider the sets of curves

$$S'_m + S'_{n+1} + \cdots + S'_{m+p} + \cdots, \qquad (m = 1, 2, \cdots).$$

C is a curve of accumulation for each of these sets, and for any given value of m the lengths of the members of the set are all less than λ_m. From Lemma 4 it follows, therefore, that the length l of C cannot exceed λ_m, and hence,

since $\lim_{m\to\infty} \lambda_m = \Lambda(\epsilon)$, we have $l \leqq \Lambda(\epsilon)$. On the other hand C is ϵ-ergodic to M, so that $l \geqq \Lambda(\epsilon)$, i. e., $l = \Lambda(\epsilon)$.

We are now in a position to prove that $\Lambda(\epsilon) > 0$ for $\epsilon < \rho$. Suppose $\Lambda(\epsilon) = 0$. The curve C constructed above is ϵ-ergodic to M and has length zero, i. e., it is a point. A circular region of radius ϵ with this point as center would enclose M. Since $\epsilon < \rho$, this contradicts Lemma 1. Hence $\Lambda(\epsilon) > 0$.

When $\epsilon \geqq \rho$, we have $\Lambda(\epsilon) = 0$ and the above curve C degenerates to a point which is the center of Γ for $\epsilon = \rho$, and can be any point in a certain point set for $\epsilon > \rho$.

The theorem just established prompts

Definition 2. A continuous, rectifiable curve of length $\Lambda(\epsilon)$ which is ϵ-ergodic to M will be termed an *ergodic curve of M*. When $\Lambda(\epsilon) = 0$, the ergodic curve of M is said to be *degenerate* or to be an *ergodic point of M*.

One readily sees that the set of ergodic points of M existing for a given value of ϵ $(\epsilon \geqq \rho)$ is closed. We complete this result by proving

Theorem 2. *The set S_ϵ of ergodic curves of M existing for a given value of ϵ $(\epsilon < \rho)$ is closed.*

Let C be an arbitrary, continuous curve of accumulation of S_ϵ. Following Lemma 4, C is seen to be rectifiable and to have a length $l \leqq \Lambda(\epsilon)$.

Consider the succession

$$S'_1, S'_2, \cdots, S'_m, \cdots,$$

of sets of curves extracted from S_ϵ by taking S'_m to be the totality of those members C'_m of S_ϵ such that C'_m, C lie in the $1/m$-neighborhood of each other. No S'_m is vacuous. The succession converges uniformly to C, so that Lemma 2 applies to prove that C is ϵ-ergodic to M. Hence $l \geqq \Lambda(\epsilon)$, i. e., $l = \Lambda(\epsilon)$, and therefore C belongs to S_ϵ.

Further information as to the nature of $\Lambda(\epsilon)$ is given in

Theorem 3. *The ergodic function $\Lambda(\epsilon)$ is a monotone non-increasing function of ϵ which is always continuous on the right.*

To prove that $\Lambda(\epsilon)$ is a monotone non-increasing function of ϵ, let $0 < \epsilon_1 < \epsilon_2$. According to Theorem 1, there is at least one ergodic curve for $\epsilon = \epsilon_1$. Since it is ϵ_1-ergodic to M, it is *a fortiori* ϵ_2-ergodic to M, and therefore $\Lambda(\epsilon_2) \leqq \Lambda(\epsilon_1)$.

We now establish that $\Lambda(\epsilon+0)=\Lambda(\epsilon)$ for $\epsilon<\rho$. The equation obviously holds for $\epsilon\geqq\rho$. Since $\Lambda(\epsilon)$ is a monotone non-increasing function of ϵ, it is sufficient to show that the inequality

$$(2) \qquad \Lambda(\epsilon+0)<\Lambda(\epsilon),$$

is impossible. Assuming (2) to hold, we let $\epsilon_1, \epsilon_2, \cdots, \epsilon_n, \cdots$ be a monotone decreasing sequence of ϵ-values lying in the open interval $(0, \rho)$ and converging to the value ϵ in (2). Consider the succession

$$(3) \qquad S_1; S_2, \cdots, S_n, \cdots$$

of sets of curves in which S_n comprises a single, well-defined curve; this curve being an ergodic curve of M for $\epsilon=\epsilon_n$ constructed according to the principles laid down in the proof of Theorem 1. The lengths of the curves in (3) are all less than λ; where λ is any number between $\Lambda(\epsilon+0)$ and $\Lambda(\epsilon)$. A finite domain contains all the curves in (3) and therefore according to Lemma 3, there is at least one well-defined, continuous, rectifiable curve of accumulation C of the succession.

From Lemma 4 it is seen that the length of C cannot exceed λ and hence is less than $\Lambda(\epsilon)$.

Now Lemma 3 asserts that a succession

$$(4) \qquad S'_1, S'_2, \cdots, S'_m, \cdots, \qquad S'_m=S_n \ (n\geqq m),$$

of sets of curves can be found which converges uniformly to C. Consider the successions

$$(5) \qquad S'_m, S'_{m+1}, \cdots, S'_{m+p}, \cdots, \qquad (m=1, 2, \cdots).$$

Each of these successions converges uniformly to C and for any given value of m the members of the succession are all ϵ_m-ergodic to M. Hence C is ϵ_m-ergodic to M and, since $\lim\limits_{m\to\infty}\epsilon_m=\epsilon$, C is ϵ-ergodic to M.

This is a contradiction; for C cannot be ϵ-ergodic to M, since its length is less than $\Lambda(\epsilon)$. Hence $\Lambda(\epsilon+0)=\Lambda(\epsilon)$.

Our last theorem deals with the structure of M and the behavior of $\Lambda(\epsilon)$ in the neighborhood of $\epsilon=0$.

THEOREM 4. *In order that $\Lambda(+0)$ be finite, it is necessary and sufficient that M lie on a continuous, rectifiable curve.*

The sufficiency of the condition is obvious.

In order to prove the necessity of the condition, suppose $\Lambda(+0)$ to be finite and let $\epsilon_1, \epsilon_2, \cdots, \epsilon_n, \cdots$ be a monotone decreasing sequence of ϵ-values lying in the open interval $(0, \rho)$ and converging to 0. Following the proof of Theorem 3, construct the succession (3) of sets of curves. The lengths of the curves in this succession cannot exceed $\Lambda(+0)$ and are therefore less than λ, where λ denotes an arbitrary positive number greater than $\Lambda(+0)$ [$\Lambda(0)$ is as yet undefined]. On extracting the succession (4) from (3) and forming the successions (5), we see that the curve of accumulation C is ϵ_m-ergodic to M, with $\lim_{m\to\infty} \epsilon_m = 0$. Hence C contains all the points of M.

REMARKS. In Theorem 4 the length of C cannot exceed λ and therefore cannot exceed $\Lambda(+0)$. On the other hand, no continuous, rectifiable curve can contain M and have a length l less than $\Lambda(+0)$. For, if such a curve exists, $\Lambda(\epsilon_n) \leqq l$ and therefore $\Lambda(+0) \leqq l$, which is a contradiction. The length of C accordingly equals $\Lambda(+0)$.

If we now define $\Lambda(0)$ as the greatest lower bound of the lengths of the continuous, rectifiable curves which contain M, we see that Theorems 1 and 3 both hold for $0 \leqq \epsilon \leqq \rho$, provided that $\Lambda(+0)$ be finite.

TRINITY COLLEGE,
HARTFORD, CONNECTICUT.

NOTE ON A RESULT OF M. H. MARTIN CONCERNING A PARTICULAR MASS RATIO IN THE RESTRICTED PROBLEM OF THREE BODIES.

By Natalie Rein.

In his paper, "On the libration points of the restricted problem of three bodies,"[1] M. H. Martin states the following theorem:

Theorem 3. *There exists in the interval $0 < \mu < 1$ one and only one value of $\mu = \mu^*$ for which the following relations hold between the functions $\rho_1(\mu)$ and $\rho_2(\mu)$:*

$$\rho_1(\mu) < \rho_2(\mu) \quad \textit{for} \quad 0 < \mu < \mu^*; \qquad \rho_1(\mu^*) = \rho_2(\mu^*); \qquad \rho_1(\mu) > \rho_2(\mu) \quad \textit{for} \quad \mu^* < \mu < 1,$$

and here $\mu^* > \frac{1}{2}$.

Here μ is the mass of one of the two finite bodies (the unit of mass being taken as the sum of the masses of the two finite bodies), and $\rho_1(\mu)$ and $\rho_2(\mu)$ are the distances of the collinear libration points L_1 and L_2 from the mass μ (the constant distance between two finite bodies being taken as the unit of distance), which are functions of μ defined uniquely by the equations:

$$\text{(a)} \quad 1 - \rho_1 - \mu - \frac{1-\mu}{(1-\rho_1)^2} + \frac{\mu}{\rho_1^2} = 0, \quad 1 + \rho_2 - \mu - \frac{1-\mu}{(1+\rho_2)^2} - \frac{\mu}{\rho_2^2} = 0.$$

See the equations (14 a) and (14 b) of Martin's paper.

The above theorem is, however, erroneous and should be formulated in its correct form as follows:

Theorem. *In the interval $0 < \mu < 1$ the functions $\rho_1(\mu)$ and $\rho_2(\mu)$ satisfy the inequality $\rho_1(\mu) < \rho_2(\mu)$.*

That is, the particular mass ratio μ^* found by Martin does not exist.

Indeed, we can write the equations (a) in the form

$$\frac{1-\mu}{(1-\rho_1)^2} + \rho_1 = 1 - \mu + \frac{\mu}{\rho_1^2}, \qquad \frac{1-\mu}{(1+\rho_2)^2} - \rho_2 = 1 - \mu - \frac{\mu}{\rho_2^2}.$$

[1] *American Journal of Mathematics*, vol. 53 (1931), p. 167.

Then $\rho_1(\mu)$ will be defined as the positive abscissa of the intersection point of the curves

$$F_1(\rho)=\frac{1-\mu}{(1-\rho)^2}+\rho, \quad f_1(\rho)=1-\mu+\frac{\mu}{\rho^2}$$

and $\rho_2(\mu)$ as the positive abscissa of the intersection point of the curves

$$F_2(\rho)=\frac{1-\mu}{(1+\rho)^2}-\rho, \quad f_2(\rho)=1-\mu-\frac{\mu}{\rho^2}.$$

Thereby, $0<\rho_1<1$, $0<\rho_2<1$ for $0<\mu<1$ according to Martin's Theorem I.[2]

$$\text{But } \frac{dF_1}{d\rho}=\frac{2(1-\mu)}{(1-\rho)^3}+1>0, \qquad \frac{dF_2}{d\rho}=-\left[\frac{2(1-\mu)}{(1+\rho)^3}+1\right]<0$$

and
$$\left|\frac{dF_1}{d\rho}\right|>\left|\frac{dF_2}{d\rho}\right| \quad \text{for} \quad 0<\rho<1.$$

This means that the intersection point of the curves F_1 and f_1 always lies to the left of the intersection point of the curves F_2 and f_2 for both the curves f_1 and f_2 are symmetrical with respect to the straight line $f=1-\mu$ and $F_1=F_2=1-\mu$ for $\rho=0$. Hence, $\rho_1<\rho_2$ for all the values of μ within the interval $0<\mu<1$.

Dr. Martin's error entered when he derived his equation (27), which defines the value $\rho^*=\rho_1=\rho_2$. In fact, he writes:

$$\rho^4(\rho^5-6\rho^3-2\rho^2+6)=0.$$

Actually, the correct form of this equation is

$$\rho^4(\rho^5-3\rho^3-\rho^2+3)=0.$$

The examination of the equation $\rho^5-3\rho^3-\rho^2+3=0$ shows immediately that it has only two positive roots, $\rho=1$ and $\rho=\sqrt{3}$, both lying outside the interval of the values of ρ_1 and ρ_2 for $0<\mu<1$.

Miss Jenny E. Rosenthal in her paper "Note on the numerical value of a particular mass ratio in the restricted problem of three bodies,"[3] has calculated from Martin's equation (27) the numerical values of ρ^* and μ^*. This equation being incorrect, the results of Miss Rosenthal's calculations could not be correct.

Moscow, U. S. S. R.

---

[2] *Ibid.*

[3] *American Journal of Mathematics*, vol. 53 (1931), p. 258.

ON THE ADDITION OF CONVEX CURVES.

By Richard Kershner.

If C denotes a convex Jordan curve in the plane let $\Omega(C)$ denote the open, bounded domain surrounded by C, so that the closed region $\bar{\Omega}(C)$ is the logical sum $\Omega(C) + C$. If A and B are two point sets in the z-plane, where $z = x + y$, their vectorial sum [1] will be denoted by $A\ (+)\ B$.

In connection with his investigation of the distribution of the values of the Riemann zeta function,[2] Bohr [3] has studied the vectorial addition of convex curves from a geometrical point of view. He has proven,[4] among other things, that if $C_1, C_2, \cdots, C_n$ are convex curves, then either there exists a convex C_E such that

$$C_1(+)\ C_2(+) \cdots (+)\ C_n = \bar{\Omega}(C_E) \tag{1}$$

or there exist two convex curves C_E and C_I such that $\Omega(C_E) \supset \Omega(C_I)$ and

$$C_1(+)\ C_2(+) \cdots (+)\ C_n = \bar{\Omega}(C_E) - \Omega(C_I) \tag{2}$$

so that the vectorial sum is either a closed bounded region or a ring shaped region bounded by two convex curves.

Following a suggestion of Wintner, Haviland [5] applied the supporting function (Stützfunktion) of Brunn and Minkowski to the study of Bohr's problem and showed that if $h_1(\theta), h_2(\theta), \cdots, h_n(\theta)$ are the supporting functions of the curves $C_1, C_2, \cdots, C_n$ and if $h_E(\theta)$ is the supporting function of C_E, then

---

[1] By the vectorial sum of two sets A and B is meant the set $z = z_a + z_b$, where $z_a \subset A$ and $z_b \subset B$. It is clear that this addition is associative and commutative.

[2] Cf. e. g. Titchmarsh, *The Zeta Function of Riemann.*

[3] H. Bohr, "Om Addition of unendelig mange konvekse Kurver," *Danske Videnskabernes Selskab* (Forhandlinger, 1913), pp. 325-366; also, for a further study cf. H. Bohr and B. Jessen, "Om Sandsynlighedsfordelinger ved Addition of konvekse Kurver," *Danske Videnskabernes Selskab, Skrifter,* (8), vol. 12, no. 3 (1929).

[4] Cf. H. Bohr, *loc. cit.*; also H. Bohr and B. Jessen, *loc. cit.* For a short presentation of the proof of this fact cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society,* vol. 38 (1935), p. 69.

[5] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics,* vol. 55 (1933), pp. 332-334.

$$h_E(\theta) = \sum_{j=1}^{n} h_j(\theta). \tag{3}$$

Using this explicit formula and Minkowski's results on mixed area, Haviland improved on the results of Bohr concerning relations between the areas of $\Omega(C_j)$ and $\Omega(C_E)$. He also derived results with regard to lengths and curvatures. Finally, Bohr and Jessen [6] have recently shown that, with the use of the supporting functions, it is possible to determine the set of those $\sigma > 1$ for which the closure of the values attained by the logarithm of the Riemann zeta function on the line σ is a convex region; for the remaining values of $\sigma > 1$ this closure is ring-shaped. They also pointed out a geometrical criterion for the existence of an inner curve in the case of arbitrary symmetric convex curves.

The object of the present note is an analytical discussion of the inner curve C_I similar to that given by Haviland for C_E. In particular there will be obtained a rule giving the supporting function $h_I(\theta)$ of C_I in terms of the supporting functions $h_j(\theta)$ of the added curves C_j in the same manner as the rule (2) gives $h_E(\theta)$. There will also be obtained results for the length of C_I, for the curvature along this curve and for the area of $\Omega(C_I)$. As an application of the analytical results concerning $h_I(\theta)$ there is given a simple mechanical construction of C_I whenever C_I exists.[7] In particular there is a geometrical criterion that C_I exist, i. e. that (2) hold. This criterion will imply the abovementioned criterion found by Bohr and Jessen in the case of symmetric curves.

We treat first the sum of two convex curves and prove

THEOREM I_0. *Let $n = 2$ so that either $C_1 \, (+) \, C_2 = \bar{\Omega}(C_E)$ or*

$$C_1 \, (+) \, C_2 = \bar{\Omega}(C_E) - \Omega(C_I).$$

It may be supposed that $l_1 \geqq l_2$ where l_j is the length of C_j. Let $\tilde{C}_2$ represent the curve obtained by rotating C_2 about the origin of the plane through an angle π. Then the curve C_I exists if and only if $\tilde{C}_2$ may be placed in $\Omega(C_1)$ by a translation.

Proof. Suppose that C_I exists so that $\Omega(C_I)$ is not empty. It is clear that a translation of C_2 effects only a translation on the region $\bar{\Omega}(C_E) - \Omega(C_I)$.

[6] H. Bohr and B. Jessen, "On the distribution of the values of the Riemann zeta function," *American Journal of Mathematics*, vol. 58 (1936), pp. 35-44.

[7] For further applications cf. R. Kershner and A. Wintner, "On the boundary of the range of values of $\zeta(s)$," *American Journal of Mathematics*, vol. 58 (1936), pp. 421-425.

Hence we may translate C_2 into a new position C_2^* in such a way that $\Omega(C_I^*)$ includes the origin of the plane, where, of course,

$$C_1(+)C_2^* = \bar{\Omega}(C_E^*) - \Omega(C_I^*).$$

Since the origin is in $\Omega(C_I^*)$, no point of C_2^* is symmetrical, with respect to the origin, to a point of C_1. Hence, the curve $\tilde{C}_2^*$, obtained by rotating C_2^* about the origin through an angle π, does not intersect C_1. Since $l_2 \leqq l_1$, either $\tilde{C}_2^*$ is in $\Omega(C_1)$ or $\bar{\Omega}(C_1)$ and $\bar{\Omega}(\tilde{C}_2^*)$ are disjoint. Now the latter case is impossible. For in this case the set $\bar{\Omega}(C_1)$ $(+)$ $\bar{\Omega}(\tilde{C}_2^*)$ would not contain the origin, whereas it is obvious [8] that

$$\bar{\Omega}(C_1)\ (+)\ \bar{\Omega}(C_2^*) = \bar{\Omega}(C_E^*)$$

and we have supposed that the origin was in $\Omega(C_I^*) \subset \Omega(C_E^*)$. Hence $\tilde{C}_2^*$ is in $\Omega(C_1)$ and the "only if" of Theorem I_0 is proven. The "if" may be shown by following the above proof in the reverse direction.

Next we prove

THEOREM II_0. *Let $n=2$ and $l_1 \geqq l_2$. Let T_I be the set of all points (x,y) of the plane which satisfy*

$$x\cos\theta + y\sin\theta < h_1(\theta) - h_2(\theta+\pi) \quad \textit{for all } \theta. \tag{4}$$

Then if T_I is not empty C_I exists and $T_I = \Omega(C_I)$; if T_I is empty C_I does not exist.

Proof. It is clear, here also, that a translation of C_2 effects a translation of T_I. Hence we may translate C_2 into a new position C_2^* in such a way that T_I^* contains the origin (where, of course, T_I^* is defined in terms of C_1 and C_2^* exactly as T_I was defined in terms of C_1 and C_2) if and only if T_I is not empty. But T_I^* contains the origin if and only if

$$0 < h_1(\theta) - h_2^*(\theta+\pi) \text{ for all } \theta \tag{5}$$

where $h_2^*(\theta)$ is the supporting function of C_2^*. Since the supporting function $h(\theta)$ of any convex curve is defined [9] as

$$h(\theta) = \max\,[x\cos\theta + y\sin\theta], \qquad (x,y) \text{ on } C, \tag{6}$$

it is clear that $h_2^*(\theta+\pi)$ is the supporting function of the curve $\tilde{C}_2^*$

[8] Cf. E. K. Haviland, *loc. cit.*, p. 332.

[9] Cf., e. g., G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis I.* (Berlin, 1925), p. 106.

obtained by rotating $C_2{}^*$ about the origin through an angle π. Hence the relation (5) implies that $\tilde{C}_2{}^*$ is within C_1. Thus T_I is not empty if and only if $\tilde{C}_2$ may be placed entirely within C_1 by means of a translation. Consequently, by Theorem I_0, T_I is empty if and only if C_I does not exist and the last part of Theorem II has been proven.

Suppose now that T_I is not empty.

First, $T_I \subset \bar{\Omega}(C_E)$. For, it is seen from (3) and (6) that $\bar{\Omega}(C_E)$ is the set of points (x, y) satisfying

$$x \cos \theta + y \sin \theta \leqq h_1(\theta) + h_2(\theta) \text{ for all } \theta; \tag{7}$$

and (4) implies (7), since $-h_2(\theta + \pi) < h_2(\theta)$. In fact, $h_2(\theta) + h_2(\theta + \pi)$ (> 0) is precisely the width of the curve C_2 in the direction θ.

Next $T_I \subset \Omega(C_I)$. (It has already been shown that C_I exists). For let $z_j = x_j + iy_j$ be a point of C_j, where $j = 1, 2$, and let $\theta_1 = \theta(z_1)$ be a value of θ such that

$$x_1 \cos \theta_1 + y_1 \sin \theta_1 = h_1(\theta_1). \tag{8}$$

Then, since z_2 is on C_2, the definition (6) of $h_2(\theta)$ gives

$$x_2 \cos \theta + y_2 \sin \theta \leqq h_2(\theta) \text{ for all } \theta; \tag{9}$$

hence, in particular (9) holds for $\theta = \theta_1 + \pi$ and

$$x_2 \cos \theta_1 + y_2 \sin \theta_1 \geqq -h_2(\theta_1 + \pi). \tag{10}$$

From (8) and (10)

$$(x_1 + x_2) \cos \theta_1 + (y_1 + y_2) \sin \theta_1 \geqq h_1(\theta_1) - h_2(\theta_1 + \pi),$$

so that $C_1 \,(+)\, C_2 = \bar{\Omega}(C_E) - \Omega(C_I) \subset \bar{\Omega}(C_E) - T_I$, i. e. $T_I \subset \Omega(C_I)$.

Finally, $T_I \supset \Omega(C_I)$. For, let B be the boundary of T_I and let $z_0 = x_0 + iy_0$ be any point of B. Then, by the definition of T_I,

$$x_0 \cos \theta + y_0 \sin \theta \leqq h_1(\theta) - h_2(\theta + \pi) \text{ for all } \theta, \tag{11}$$

while, for at least one θ, say $\theta = \theta_0$,

$$x_0 \cos \theta_0 + y_0 \sin \theta_0 = h_1(\theta_0) - h_2(\theta_0 + \pi). \tag{12}$$

Let $z_2 = z_2(\theta_0) = x_2(\theta_0) + iy_2(\theta_0)$ be a point of C_2 in the direction $\theta_0 + \pi$, i. e. satisfying

$$x_2 \cos(\theta_0 + \pi) + y_2 \sin(\theta_0 + \pi) = h_2(\theta_0 + \pi). \tag{13}$$

From (6),

(14) $$x_2 \cos(\theta + \pi) + y_2 \sin(\theta + \pi) \leqq h_2(\theta + \pi) \text{ for all } \theta.$$

Let $z_0 - z_2 = z_1 = x_1 + iy_1$. Then, from (12), (13) and (11), (14)

$$x_1 \cos\theta_0 + y_1 \sin\theta_0 = h_1(\theta_0) \text{ and } x_1 \cos\theta + y_1 \sin\theta \leqq h_1(\theta) \text{ for all } \theta.$$

Thus z_1 is a point of C_1, as well as z_2 a point of C_2, and $z_1 + z_2 = z_0$ where z_0 was any point of B. Thus, $B \subset \bar{\Omega}(C_E) - \Omega(C_I)$. Since T_I is obviously convex, this implies $T_I \supset \Omega(C_I)$. But it has already been shown that $T_I \subset \Omega(C_I)$ so that the proof of Theorem II_0 is complete.

Theorem II_0 implies a simple mechanical way in which $\Omega(C_I)$ may be determined: $\Omega(C_I)$ is, up to a translation, the locus of a point which is considered as rigidly attached to C_2 when $\bar{C}_2$ is translated in all possible manners in $\Omega(C_1)$. This fact, which is completely equivalent to Theorem II_0, merely expresses the amount of indeterminateness in the choice of the translation which occurred in the proof of Theorem I_0. It would be possible to give an entirely geometrical proof of Theorem II_0 based on this reasoning.

Since $B = C_I$, every point z_0 of C_I satisfies (11) and (12). From these inequalities it does not follow that the supporting function $h_1(\theta)$ of C_I is identically $h_1(\theta) - h_2(\theta + \pi)$. However (11) gives immediately

$$h_I(\theta) \leqq h_1(\theta) - h_2(\theta + \pi) \text{ for all } \theta.$$

Furthermore, if there is a unique supporting line through a given point z_0 of C_I, i. e. if

$$x_0 \cos\theta + y_0 \sin\theta = h_I(\theta)$$

is satisfied for a unique $\theta = \theta_0$, then by (12)

$$h_I(\theta_0) = h_1(\theta_0) - h_2(\theta_0 + \pi).$$

The curve C_I is said to have a corner at $z = z_0$ if the supporting line through z_0 is not unique. The above remarks imply

THEOREM III_0. *The supporting function $h_I(\theta)$ of C_I satisfies*

(15) $$h_I(\theta) \leqq h_1(\theta) - h_2(\theta + \pi) \text{ for all } \theta$$

and the equality sign holds save at most for the θ-intervals corresponding to corners of C_I.

We now give the corresponding theorems for the sum of n convex curves.

THEOREM I. *Let $C_1, C_2, \cdots, C_n$ be n convex curves of lengths $l_1, l_2, \cdots, l_n$ respectively and call $C_1 (+) C_2 (+) \cdots (+) C_n = \bar{\Omega}(C_E) - \Omega(C_I)$ where $\Omega(C_I)$ represents the empty set if C_I does not exist. Suppose that $l_1 \geqq l_2 \cdots \geqq l_n$. Call $C_2 (+) C_3 (+) \cdots (+) C_n = \bar{\Omega}(D_E) - \Omega(D_I)$ where $\Omega(D_I)$ is the empty set if D_I does not exist. Then C_I exists if and only if $\tilde{D}_E$, the curve obtained by rotating D_E about the origin through an angle π, can be placed in $\Omega(C_1)$ by a translation.*

THEOREM II. *Let T_I be the set of all points z satisfying*

$$x \cos \theta + y \sin \theta < h_1(\theta) - \sum_{j=2}^{n} h_j(\theta + \pi) \quad \text{for all } \theta, \tag{16}$$

where $h_j(\theta)$ is the supporting function of C_j $(j = 1, 2, \cdots, n)$. Then if T_I is not empty C_I exists and $T_I = \Omega(C_I)$; if T_I is empty C_I does not exist.

Proof. Since the supporting function $g_E(\theta)$ of D_E satisfies, by (3), the relation

$$g_E(\theta) = \sum_{j=2}^{n} h_j(\theta), \tag{17}$$

it may be shown, as it was in the case of two curves, that Theorem I is implied by Theorem II; so it is sufficient to prove the latter.

Now Theorem II is true for the case of two curves. Suppose Theorem II has been proven for the case of $n - 1$ curves. Then, using the notations of Theorem I and calling $g_I(\theta)$ the supporting function of D_I, we have

$$g_I(\theta) \leqq h_2(\theta) - \sum_{j=3}^{n} h_j(\theta + \pi). \tag{18}$$

Now it is well known that if $h(\theta)$ is the supporting function and l the length of a convex curve,

$$l = \int_0^{2\pi} h(\phi)\, d(\phi). \tag{19}$$

Let λ_I be the length of D_I. Then, by (18) and (19),

$$\lambda_I \leqq l_2 - \sum_{j=3}^{n} l_j. \tag{20}$$

Suppose first that T_I is empty. Then, as in Theorem II$_0$, the curve $\tilde{D}_E$ cannot be placed within C_1 by a translation. On the other hand, the length λ_I of D_I is, by (20), less than l_2, and, since $l_1 \geqq l_2$, less than l_1. Hence, there exists a

convex curve E between D_I and D_E, i. e. in $\bar{\Omega}(D_E) - \Omega(D_I)$, of length less than or equal to l_1 such that $\bar{E}$ cannot be placed within C_1 by a translation. The region $C_1 \, (+) \, E$ is, by Theorem I_0, a closed convex region, and a fortiori, C_I does not exist.

Suppose now that T_I is not empty. Then $C_1 \, (+) \, D_E = \bar{\Omega}(C_E) - T_I$ by (17) and Theorem II_0. Let F be any convex curve in $\bar{\Omega}(D_E) - \Omega(D_I)$ and let $k(\theta)$ be its supporting function. Let $C_1 \, (+) \, F = \bar{\Omega}(G_E) - \Omega(G_I)$ where $\Omega(G_I)$ represents the empty set if G_I does not exist. Then the supporting function of G_E is, by (3), equal to $h_1(\theta) + k(\theta)$ and the supporting function of G_I is, by Theorem II_0, equal to $h_1(\theta) - k(\theta + \pi)$, except, at most, at the possible corners of G_I.

Since $k(\theta) \leqq g_E(\theta) = \sum_{j=2}^{n} h_j(\theta)$ for all θ, we have

$$h_1(\theta) - g_E(\theta + \pi) \leqq h_1(\theta) - k(\theta + \pi)$$
$$\leqq h_1(\theta) + k(\theta) \leqq h_1(\theta) + g_E(\theta) \quad \text{for all } \theta.$$

Hence $\bar{\Omega}(G_E) - \Omega(G_I) \subset \Omega(C_E) - T_I$. But F was any convex curve in $\bar{\Omega}(D_E) - \Omega(D_I)$. Consequently

$$C_1 \, (+) \, D_E = C_1 \, (+) \, C_2 \, (+) \cdots (+) \, C_n \tag{21}$$

and $T_I = \Omega(C_I)$.

Relations (17) and (21), together with Theorem III_0, give immediately

THEOREM III. *The supporting function $h_I(\theta)$ of C_I satisfies*

$$h_I(\theta) \leqq h_1(\theta) - \sum_{j=2}^{n} h_j(\theta + \pi), \text{ for all } \theta, \tag{22}$$

and the equality sign holds save at most for the θ-intervals corresponding to corners of C_I. Thus, if $\{z_k\}$ is the sequence of corners (if any) of C_I and $\{\Theta_k\}$ is the sequence of the corresponding open θ-intervals and if Θ denotes the closed set $\Theta = [0, 2\pi] - \Sigma\Theta_k$ then

$$h_I(\theta) = h_1(\theta) - \sum_{j=2}^{n} h_j(\theta + \pi), \text{ if } \theta \subset \Theta$$
$$h_I(\theta) = x_j \cos\theta + y_j \sin\theta, \text{ if } \theta \subset \Theta_j; \; z_j = x_j + iy_j.$$

It may be mentioned that all the above results are true, with appropriate changes in statement, for the vectorial sum of convex bodies of any number of dimensions.

We now proceed to the length, area and curvature relations implied by (22).

Formula (19), for the length of a convex curve, together with (22), gives immediately

$$l_I \leqq l_1 - \sum_{j=2}^{n} l_j \tag{23}$$

where l_I is the length of C_I.

It is well known that the area $A(C)$ of $\Omega(C)$, where C is a convex curve with supporting function $h(\theta)$, is represented by the formula

$$A(C) = \tfrac{1}{2}\int_0^{2\pi} [h^2(\phi) - (h'(\phi))^2] d\phi. \tag{24}$$

Let $M(C, D)$, where C and D are convex curves, represent the Minkowski mixed area [10] of the regions $\Omega(C)$ and $\Omega(D)$. Then by (22), (24), and (17)

$$A(C_I) \leqq \tfrac{1}{2}\int_0^{2\pi} [(h_1(\phi) - g_E(\phi+\pi))^2 - (h'_1(\phi) - g'_E(\phi+\pi))^2] d\phi$$

or

$$A(C_I) \leqq \tfrac{1}{2}\int_0^{2\pi} [h_1^2(\phi) - (h'_1(\phi))^2] d\phi + \tfrac{1}{2}\int_0^{2\pi} [g_E^2(\phi+\pi) - (g'_E(\phi+\pi))^2] d\phi$$
$$- \int_0^{2\pi} [h_1(\phi) g_E(\phi+\pi) - h'_1(\phi) g'_E(\phi+\pi)] d\phi$$

i. e.

$$A(C_I) \leqq A(C_1) + A(D_E) - 2M(C_1, D_E). \tag{25}$$

Since, according to Minkowski,

$$M(C, D) \geqq \sqrt{A(C)A(D)},$$

(25) becomes

$$\sqrt{A(C_I)} \leqq \sqrt{A(C_1)} - \sqrt{A(D_E)}$$

or, since [11] $\sqrt{A(D_E)} \geqq \sum_{j=2}^{n} \sqrt{A(C_j)}$,

$$\sqrt{A(C_I)} \leqq \sqrt{A(C_1)} - \sum_{j=2}^{n} \sqrt{A(C_j)}.$$

If (25) is compared with the corresponding formula of Haviland,

$$A(C_E) = A(C_1) + A(D_E) + 2M(C_1, D_E),$$

one finds

$$A = A(C_E) - A(C_I) \geqq 4M(C_1, D_E) = 4\sum_{j=2}^{n} M(C_1, C_j),$$

[10] Cf. e. g., W. Blaschke, *Kreis und Kugel* (Leipzig, 1916), pp. 106-107. or T. Bonnesen, *Les Problèmes des Isoperimètres* (Paris, 1929), Ch. V.

[11] Cf. E. K. Haviland, *loc. cit.*, p. 334.

where A is the area of the ring-shaped region (2). Hence

$$A \geqq 4\sqrt{A(C_1)} \sum_{j=2}^{n} \sqrt{A(C_j)}.$$

We summarize the above results on area and length relations in

THEOREM IV. *Let l_j be the length of C_j and $A(C_j)$ the area of $\Omega(C_j)$. Let l_I be the length of C_I and $A(C_I)$ the area of $\Omega(C_I)$. Let A be the area of $C_1 (+) C_2 (+) \cdots (+) C_n$. Then, if C_I exists,*

(i) $$l_I \leqq l_1 - \sum_{j=2}^{n} l_j$$

and

(ii) $$\sqrt{A(C_I)} \leqq \sqrt{A(C_1)} - \sum_{j=2}^{n} \sqrt{A(C_j)}.$$

Whether C_I exists or not

(iii) $$A \geqq 4 \sum_{j=2}^{n} M(C_1, C_j),$$

where $M(C_1, C_j)$ is the mixed area of $\Omega(C_1)$ and $\Omega(C_j)$, and

(iv) $$A \geqq 4\sqrt{A(C_1)} \sum_{j=2}^{n} \sqrt{A(C_j)}.$$

In (i) *and* (iii) *the equality sign holds if C_I has no corners.*

Suppose, now, that the supporting functions $h_j(\theta)$ have continuous second derivatives, and that the radius of curvature $\rho_j(\theta)$ of C_j is defined, so that

(26) $$\rho_j(\theta) = h_j(\theta) - h_j''(\theta).$$

Suppose that the point in the direction θ_1 on C_I is not a corner or a cluster point of corners, so that

$$h_I(\theta) = h_1(\theta) - \sum_{j=2}^{n} h_j(\theta + \pi), \text{ if } \theta_1 - \epsilon \leqq \theta \leqq \theta_1 + \epsilon$$

for some $\epsilon > 0$. Then, by (26),

$$\rho_I(\theta_1) = \rho_1(\theta_1) - \sum_{j=2}^{n} \rho_j(\theta_1 + \pi)$$

where $\rho_I(\theta_1)$ is the radius of curvature of C_I at the point in the direction θ_1. In particular, if C_I has no corners then $\rho_I(\theta) \geqq \sum_{j=2}^{n} \rho_j(\theta + \pi)$ for all θ where

the equality sign holds at most for discrete θ-values. It is easy to see, in view of Theorem II, that the converse of this last statement is also true. Thus we have

THEOREM V. *If $h_j(\theta)$ has a continuous second derivative for $j = 1, 2, \cdots, n$, then a necessary and sufficient condition that C_I exist and have no corners is that $\rho_1(\theta) \geqq \sum_{j=2}^{n} \rho_j(\theta + \pi)$ for all θ where the equality sign holds at most for discrete θ-values. In general, if the point in the direction θ_1 on C_I is not a corner or a cluster point of corners of C_I then the radius of curvature $\rho_I(\theta)$ of C_I is defined for $\theta = \theta_1$, and*

$$\rho_I(\theta_1) = \rho_1(\theta_1) - \sum_{j=2}^{n} \rho_j(\theta_1 + \pi).$$

THE JOHNS HOPKINS UNIVERSITY.

THE IDEMPOTENT AND NILPOTENT ELEMENTS OF A MATRIX.

By JOHN WILLIAMSON.

Introduction. Associated with each latent root or characteristic number of a square matrix A are two matrices, the principal idempotent and nilpotent elements. These two matrices are polynomials in A and are uniquely determined by certain simple conditions.[1] In the following pages we show how analogous results hold, when all operations are restricted to the field K, in which the elements of A lie. Since, as a rule, the characteristic numbers of A do not lie in the field K, we replace them, or more exactly, their corresponding linear factors in the characteristic equation of A, by the irreducible factors $p_i(x)$ of the characteristic equation of A. We prove that, associated with each *characteristic divisor* $p_i(x)$, there are two matrices which are polynomials in A with coefficients in K and which may justifiably be called the principal idempotent and nilpotent elements of A associated with $p_i(x)$. In case $p_i(x) = x - x_i$ is linear, these two matrices are the principal idempotent and nilpotent elements associated with the characteristic number x_i. These idempotent and nilpotent elements assume a very simple form, when the matrix A is reduced to a suitable canonical form. This canonical form is analogous to the classical canonical form of a matrix.[2] In the final section we use this canonical form to deduce theorems on matrices with elements in a field K from known theorems on matrices with elements in an algebraically closed field.

1. *Preliminary lemmas.* Let K be a commutative field of characteristic zero and let $p(x)$ be a polynomial in the ring $K[x]$ of degree ν, with leading coefficient unity, and irreducible in $K[x]$.

LEMMA 1. *For every positive integer m there exists a unique polynomial $g(x)$ in $K[x]$ of degree less than $m\nu$ and such that*

$$(x - g(x))^m \equiv 0 \mod [p(x)]^m, \tag{1}$$

$$(x - g(x))^i \not\equiv 0 \mod [p(x)]^m, \quad 0 \leqq i < m, \tag{2}$$

and

$$p[g(x)] \equiv 0 \mod [p(x)]^m. \tag{3}$$

[1] J. H. Wedderburn, "Lectures on matrices," *American Mathematical Colloquium Publications* (1934), pp. 29 and 30.

[2] Wedderburn, *op. cit.*, p. 123.

We shall prove this lemma by induction on m. If $g(x)$ satisfies (1),

$$(4) \qquad g(x) \equiv x \mod p(x),$$

and, since the only polynomial of degree less than ν satisfying (4) is x itself, our lemma is true, when $m = 1$, with $g(x) = x$. Since $p(x)$ is irreducible, $p(x)$ and its derivative $p'(x)$ are relatively prime. Accordingly there exist two uniquely determined polynomials in $K[x]$, $h(x)$ of degree less than ν and $r(x)$ of degree less than $\nu - 1$, such that

$$(5) \qquad h(x)p'(x) + r(x)p(x) = 1.$$

If,

$$(6) \qquad g_m(x) = x + h_1 p + h_2 p^2 + \cdots + h_{m-1}p^{m-1}, \quad m \geqq 2,$$

where $p = p(x)$ and each $h_i = h_i(x)$ is a polynomial in $K[x]$ of degree less than ν, then $g_m(x)$ is the most general polynomial in $K[x]$ of degree less than $m\nu$, which satisfies (4). Further,

$$(7) \qquad p(g_m) = p + p'[h_1 p + \cdots + h_{m-1}p^{m-1}] \\ + \cdots + p^{(\nu)}/\nu!\, [h_1 p + h_2 p^2 + \cdots + h_{m-1}p^{m-1}]^{\nu}.$$

When the right hand of (7) is expanded in powers of p, it is apparent that h_{m-1} occurs only in terms containing a factor p^s where $s \geqq m - 1$. Hence, if we consider the value of $p(g_m)$ modulo p^{m-1}, all terms involving h_{m-1} disappear. Accordingly,

$$(8) \qquad p(g_m) \equiv p(g_{m-1}) \mod p^{m-1},$$

where $g_{m-1}(x)$ is obtained from $g_m(x)$ in (6) by putting $h_{m-1} = 0$. It is therefore an immediate consequence of (8) that, if

$$p(g_m) \equiv 0 \mod p^m,$$
$$p(g_{m-1}) \equiv 0 \mod p^{m-1}.$$

Let us now make the induction assumption, that $g_{m-1}(x)$ is the unique polynomial of Lemma 1, satisfying (1), (2) and (3), when m is replaced by $m - 1$. Hence

$$(9) \qquad p(g_{m-1}) \equiv k_{m-1}p^{m-1} \mod p^m,$$

where k_{m-1} is a uniquely determined polynomial of degree less than ν. From (7) and (8) we see that

$$p(g_m) \equiv p(g_{m-1}) + p'h_{m-1}p^{m-1} \mod p^m,$$
$$\equiv (k_{m-1} + p'h_{m-1})p^{m-1} \mod p^m \text{ by } (9).$$

Therefore,

$$p(g_m) \equiv 0 \mod p^m,$$

if and only if,

(10) $$k_{m-1} + p'h_{m-1} \equiv 0 \bmod p.$$

But, it follows from (5) that

$$k_{m-1} + p'h_{m-1} \equiv (hk_{m-1} + k_{m-1})p' \bmod p.$$

Since p' is relatively prime to p, (10) is true, if and only if,

$$h_{m-1} + hk_{m-1} \equiv 0 \bmod p.$$

This last congruence determines k_{m-1} and therefore $g_m(x)$ uniquely. Since, when $m = 2$, $k_{m-1} = 1$, $h_1 = -h \neq 0$. Hence $g_m(x)$ satisfies (2) and $g(x) = g_m(x)$ is the unique polynomial of Lemma 1 satisfying (1), (2) and (3).

If $\nu = 1$, so that $p(x) = x - a$, $g_m(x)$ is the same for all values of $m \geqq 2$. In fact $g_m(x) = a$; for $a = x - (x - a)$ and $a - a = 0 \equiv 0 \bmod (x-a)^m$ for all positive integral values of m. This however is not true in general.

Let P be a square matrix of order ν with elements in K, whose characteristic polynomial is the irreducible polynomial $p(x)$. For definiteness we may take P to be the companion matrix of $p(x)$.[3] There is a one to one correspondence between the matrices $f(P)$, where $f(x)$ is a polynomial of $K[x]$, and the polynomials $f(\theta)$, where θ is a zero of $p(x)$. Since the correspondence is preserved under addition and multiplication it is apparent that the totality of matrices $f(P)$ forms a field simply isomorphic with the algebraic extension field $K(\theta)$. Consequently the minimal or reduced characteristic polynomial of $f(P)$ is an irreducible polynomial $q(x)$ of degree μ where $\nu/\mu = \sigma$, an integer.[4] Accordingly the characteristic polynomial of $f(P)$ is $[q(x)]^\sigma$ and the elementary factors of $f(P)$ are $q(x)$ counted σ times.[5] If Q is the companion matrix of $q(x)$, $f(P)$ is therefore similar in K to the *diagonal block matrix*

$$Q_\sigma = [Q, Q, \cdots, Q],$$

consisting of the matrix Q repeated σ times in the leading diagonal; that is, there exists a non-singular matrix C with elements in K, such that

(11) $$Cf(P)C^{-1} = Q_\sigma.$$

Let E_i denote the unit matrix of order i and U_i the auxiliary unit matrix of order i, that is the matrix all of whose elements are zero except those in the

[3] C. C. MacDuffee, *The Theory of Matrices*, Berlin, 1933, p. 20.

[4] B. L. Van der Waerden, *Moderne Algebra*, vol. 1, p. 98.

[5] The powers of the distinct irreducible factors of the invariant factors of $A - xE$ are called the elementary factors of A.

diagonal to the right of the leading one, each of which is unity. Then the matrix W of order k

$$(12) \qquad W = \begin{pmatrix} f(P) & E_\nu & 0 & \cdots & 0 \\ 0 & f(P) & E_\nu & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & E_\nu \\ 0 & 0 & 0 & \cdots & f(P) \end{pmatrix}$$

may be written in the convenient form

$$(13) \qquad W = f(P) \cdot E_k + E_\nu \cdot U_k.$$

It now follows from (11) that

$$C \cdot E_k W (C \cdot E_k)^{-1} = Q_\sigma \cdot E_k + E_\nu \cdot U_k.$$

By a simple rearrangement of the rows and the same rearrangement of the columns, this last matrix may be reduced to a diagonal block matrix of σ blocks each block being the matrix $Q \cdot E_k + E_\mu \cdot U_k$. Since the minimal equation of $Q \cdot E_k + E_\mu \cdot U_k$ is $[q(x)]^k = 0$,[6] we have proved;

LEMMA 2. *The elementary factors of the matrix W defined by* (12) *are $[q(x)]^k$ repeated σ times, where $q(x)$ is the minimal polynomial of $f(P)$ and is of degree $\mu = \nu/\sigma$.*

2. *Matrices with a single characteristic divisor.* Let A be a square matrix with elements in K, whose minimal equation is $\phi(x) = [p(x)]^m = 0$, where $p(x)$ is the irreducible polynomial of the previous section, and let $g(x)$ be the unique polynomial of Lemma 1 satisfying (1), (2) and (3). If $g(x) = y$, as a consequence of (1) and (2),

$$(14) \qquad (x-y)^m \equiv 0, \quad (x-y)^i \not\equiv 0 \bmod [p(x)]^m, \quad 0 \leqq i < m.$$

Hence, if $f(x)$ is any polynomial of $K[x]$,

$$(15) \quad f(x) \equiv f(y) + f'(y)(x-y) + \cdots + 1/(m-1)!\, f^{(m-1)}(y)(x-y)^{m-1} \bmod [p(x)]^m.$$

If we substitute the matrix A for the indeterminate x, we deduce from (3) that, if $g(A) = A_1$,

$$(16) \qquad p(A_1) = 0$$

and, from (14), that the matrix $\eta_1 = A - A_1$ satisfies

$$\eta_1^m = 0, \quad \eta_1^i \neq 0, \quad 0 \leqq i < m.$$

[6] Wedderburn, *op. cit.*, p. 124.

Further, from (15), we have

(17) $f(A) = f(A_1) + f'(A_1)\eta_1 + \cdots + 1/(m-1)!\, f^{(m-1)}(A_1)\eta_1^{m-1}$.

In (17) the polynomials $f(A_1), f'(A_1)$ etc. may all be reduced modulo $p(A_1)$ by virtue of (16).

We now prove the theorem.

THEOREM I. *If $q(x) = 0$ is the minimal equation of $f(A_1)$,*

$$\textit{and } f'(A_1) = f''(A_1) = \cdots = f^{(s-1)}(A_1) = 0, \quad f^{(s)}(A_1) \neq 0,$$

then the minimal equation of $f(A)$ is $[q(x)]^\kappa = 0$, where κ is a positive integer uniquely defined by the inequalities,

$$(\kappa - 1)s < m \leqq \kappa s. \tag{18}$$

Proof. Since $q[f(A_1)] = 0$, $q[f(A)]^\kappa = 0$, while, since $f^{(s)}(A_1) \neq 0$, $q[f(A)]^{\kappa-1} \neq 0$. But $q(x)$ is irreducible[7] and so the minimal equation of $f(A)$ is $[q(x)]^\kappa = 0$.

Moreover

$$q[f(A)] \equiv q[f(A_1)] + q'[f(A_1)] \frac{f^{(s)}(A_1)}{s!} \eta_1^s \mod \eta_1^{s+1}$$

and

$q'[f(A_1)] \dfrac{f^{(s)}(A_1)}{s!} \neq 0$. Hence, as a consequence of the definition of η_1,

$$q[f(x)] \equiv 0 \mod [p(x)]^s, \quad q[f(x)] \not\equiv 0 \mod [p(x)]^{s+1}. \tag{19}$$

We have therefore the corollary. *The integer s may be determined by the congruences* (19).

3. *Matrices with more than one characteristic divisor.* Let A be a square matrix with elements in K, whose reduced characteristic equation is

$$\phi(x) = \prod_{i=1}^{t} [p_i(x)]^{m_i} = 0$$

where $p_i(x)$ $(i = 1, 2, \cdots, t)$, are t distinct irreducible polynomials in $K[x]$ of degrees ν_i respectively each with leading coefficient unity. The degree of $\phi(x)$ is accordingly given by the equation,

$$\nu = \sum_{i=1}^{t} \nu_i m_i.$$

If

$$h_i(x) = \phi(x)/[p_i(x)]^{m_i},$$

[7] The totality of matrices $f(A_1)$ is simply isomorphic to the field $K(\theta)$.

$h_i(x)$ and $[p_i(x)]^{m_i}$ are relatively prime, so that there exist two polynomials $M_i(x)$ and $N_i(x)$ in $K[x]$, of degrees less than $m_i\nu_i$ and $\nu - m_i\nu_i$ respectively, such that

$$M_ih_i + p_i^{m_i}N_i = 1.$$

Since, when $i \neq j$, $h_i \equiv 0 \bmod p_j^{m_j}$, it follows that

$$\sum_{i=1}^{t} M_ih_i \equiv 1 \bmod p_j^{m_j} \qquad (j = 1, 2, \cdots, t),$$

and accordingly that

$$\sum_{i=1}^{t} M_ih_i \equiv 1 \bmod \phi(x). \tag{20}$$

Since the congruence (20) is of degree less than ν in x, it must be an identity and on writing $M_ih_i = \phi_i$ we have

$$\sum_{i=1}^{t} \phi_i = 1. \tag{21}$$

It is an immediate consequence of the definition of ϕ_i that

$$\phi_i\phi_j \equiv 0 \bmod \phi(x) \qquad (i \neq j), \tag{22}$$

and of (21) and (22) that

$$\phi_i^2 \equiv \phi_i \not\equiv 0 \bmod \phi(x). \tag{23}$$

Let $g_i(x) = y_i$ be the polynomial of Lemma 1, when $p(x)$, ν and m are replaced by $p_i(x)$, ν_i and m_i respectively. Then

$$[(x - y_i)\phi_i]^{m_i} \equiv (x - y_i)^{m_i}\phi_i \equiv 0 \bmod \phi(x),$$
$$[(x - y_i)\phi_i]^{j} \not\equiv 0 \bmod \phi(x), \quad 0 \leqq j < m_i.$$

Consequently, if $f(x)$ is any polynomial of $K[x]$,

$$\begin{aligned}
f(x) &= f(x)\sum_{i=1}^{t}\phi_i = \sum_{i=1}^{t} f(x)\phi_i \\
&\equiv \sum_{i=1}^{t} \{f(y_i)\phi_i + f'(y_i)(x - y_i)\phi_i \\
&\qquad + \cdots + 1/(m_i - 1)!\, f(y_i)^{(m_i-1)}(x - y_i)^{m_i-1}\phi_i\} \bmod \phi(x), \\
&\equiv \sum_{i=1}^{t} \{f(y_i\phi_i)\phi_i + f'(y_i\phi_i)(x - y_i)\phi_i \\
&\qquad + \cdots + 1/(m_i - 1)!\, f(y_i\phi_i)^{(m_i-1)}(x - y_i)^{m_i-1}\phi_i\} \bmod \phi(x).
\end{aligned}$$

We now replace the indeterminate x by the matrix A and write

$$\phi_i(A) = \phi_i, \quad A_i = g_i(A)\phi_i(A), \quad (A - A_i)\phi_i = \eta_i,$$

so that

$$(24)\quad \sum_{i=1}^{t}\phi_i = E, \text{ the unit matrix, } \phi_i{}^2 = \phi_i \neq 0,\ \phi_i\phi_j = 0,\ i \neq j,\ \eta_i{}^{m_i} = 0,$$
$$\eta_i{}^j \neq 0,\ 0 \leqq j < m_i \qquad (i = 1, 2, \cdots, t).$$

Further, since

$$p_i(y_i) \equiv 0 \mod p_i{}^{m_i},$$
$$p_i(y_i)\phi_i \equiv p_i(y_i\phi_i)\phi_i \equiv 0 \mod \phi(x),$$

so that

$$(25)\qquad p_i(A_i)\phi_i = 0 \qquad (i = 1, 2, \cdots, t).$$

If $\phi_i = E$, that is, if $t = 1$, the minimal equation of A_i is $p_i(x) = 0$. Otherwise the minimal equation of A_i is $xp(x) = 0$, since, by (25), $A_ip_i(A_i) = 0$ and $p_i(x)$ is irreducible. If $p_i(x)$ is linear, so that $p_i(x) = x - a_i$, $A_i = A_i\phi_i = a_i\phi_i$; but as a rule it is simpler in this case to take $A_i = a_iE$.

It is apparent that the matrices ϕ_i and η_i, which satisfy (24), are respectively idempotent and nilpotent and that

$$A = A_1\phi_1 + \eta_1 + A_2\phi_2 + \eta_2 + \cdots + A_t\phi_t + \eta_t.$$

More generally any polynomial $f(A)$ may be written in the form

$$(26)\qquad f(A) = \sum_{i=1}^{t}\{f(A_i)\phi_i + f'(A_i)\eta_i + \cdots + 1/(m_i - 1)!\, f^{(m_i-1)}(A_i)\eta_i{}^{m_i-1}\}.$$

In (26), since η_i contains the factor ϕ_i, by virtue of (25) each polynomial $f^{(k)}(A_i)$ may be reduced modulo $p_i(A_i)$. The minimal equation of $f(A)$ may now be determined. Let $f^{(s_i)}(A_i)$ be the first of the derivatives of $f(A_i)$, which does not vanish, and let κ_i be determined by the inequalities

$$(27)\qquad (\kappa_i - 1)s_i < m_i \leqq \kappa_i s_i.$$

If $q_i(x) = 0$, is the minimal equation of $f(A_i)$, it may happen that, for two or more distinct values of i, the polynomials $q_i(x)$ coincide. If so, we write k_i for the largest of the corresponding integers κ_i, while, if $q_i(x)$ arises from only one value of i, we write $k_i = \kappa_i$. Then, *the minimal equation of* $f(A)$ *is*

$$(28)\qquad \Pi[q_i(x)]^{k_i} = 0,$$

where the product extends over all distinct polynomials $q_i(x)$. As in the corollary to Theorem 1 we see that the integers s_i in (27) may be determined from the congruences

$$q_i[f(x)] \equiv 0 \mod p_i{}^{s_i}, \qquad q_i[f(x)] \not\equiv 0 \mod p_i{}^{s_i+1}.$$

We shall call the matrices ϕ_i and η_i *the principal idempotent* and *nilpotent elements* of A associated with the characteristic divisor $p_i(x)$. The matrices ϕ_i, η_i and A_i are uniquely defined by (24) and (25) as is shown by the following theorem.

THEOREM 2. *If ψ_i and B_i are $2t$ matrices with elements in K all commutative with A, which satisfy the conditions,*

(i) $\quad B_i = B_i\psi_i = \psi_i B_i,\qquad$ (ii) $\quad p_i(B_i)\psi_i = 0,$

(iii) $\quad \sum_{i=1}^{t}\psi_i = E,\quad \psi_i^2 = \psi_i \neq 0,\qquad$ (iv) $\quad (A - B_i)\psi_i$ is nilpotent,

then $\psi_i = \phi_i$ and $B_i = A_i$ $(i = 1, 2, \cdots, t)$.

The proof of this theorem is similar to that of Wedderburn for the simpler case in which K is algebraically closed.[8] Let

$$\theta_{ij} = \phi_i\psi_j \qquad (i, j = 1, 2, \cdots, t).$$

Then, since ϕ_i is a polynomial in A and ψ_j is commutative with A,

$$\theta_{ij} = \psi_j\phi_i.$$

Further, the matrices

$$\eta_i = (A - A_i)\phi_i \quad \text{and} \quad \xi_j = (A - B_j)\psi_j.$$

are both nilpotent and are commutative. Since

$$\begin{aligned} A\theta_{ij} &= [A_i + (A - A_i)]\phi_i\psi_j = A_i\theta_{ij} + \eta_i\psi_j, \\ &= [B_j + (A - B_j)]\psi_j\phi_i = B_j\theta_{ij} + \xi_j\phi_i, \end{aligned}$$

$$(A_i - B_j)\theta_{ij} = \xi_j\phi_i - \eta_i\psi_j. \tag{29}$$

Since ξ_j and η_i are both nilpotent and all matrices on the right of (29) are commutative, $(A_i - B_j)\theta_{ij}$ is nilpotent. Consequently $[p_i(A_i) - p_i(B_j)]\theta_{ij}$ is nilpotent and by (25) $p_i(B_j)\theta_{ij}$ is nilpotent, so that for some integer k,

$$[p_i(B_j)]^k\theta_{ij} = 0. \tag{30}$$

Further as a consequence of (ii)

$$p_j(B_j)\theta_{ij} = 0. \tag{31}$$

---

[8] Wedderburn, *op. cit.*, p. 29. If K is algebraically closed the conditions are simplified since $B_i = a_i\psi_i$ and (i) and (ii) are automatically satisfied.

Since, when $i \neq j$, $[p_i(x)]^k$ and $p_j(x)$ are relatively prime there exist two polynomials $h(x)$ and $r(x)$ in $K[x]$ such that

$$h(x)[p_i(x)]^k + r(x)p_j(x) = 1.$$

Accordingly from (30) and (31) we have

$$[h(B_j)[p_i(B_j)]^k + r(B_j)p_j(B_j)]\theta_{ij} = \theta_{ij} = 0, \quad i \neq j.$$

Hence

$$\phi_i\psi_j = \psi_j\phi_i = 0, \quad i \neq j.$$

But

$$\phi_i = \phi_i E = \phi_i \sum_{j=1}^{t} \psi_j = \sum_{j=1}^{t} \phi_i\psi_j = \phi_i\psi_i = \sum_{j=1}^{t} \phi_j\psi_i = E\psi_i = \psi_i,$$

so that the first part of the theorem is proved.

As $(A - A_i)\phi_i$ and $(A - B_i)\psi_i$ are nilpotent, $(B_i - A_i)\phi_i = \rho$ is nilpotent. If $\rho \neq 0$, let $\rho^{k+1} = 0$, $\rho^k \neq 0$. Since,

$$B_i\phi_i = A_i\phi_i + \rho,$$
$$p_i(B_i)\phi_i \equiv p_i(A_i)\phi_i + p'_i(A_i)\phi_i\rho \mod \rho^2,$$

and by (ii) and (25)

(32) $$p'_i(A_i)\phi_i\rho \equiv 0 \mod \rho^2.$$

If $r_i(x)$ is the polynomial $r(x)$ of (5), when $p(x)$ is replaced by $p_i(x)$,

$$r_i(A_i)p'_i(A_i)\phi_i = \phi_i,$$

so that by (32)

$$\phi_i\rho = \rho \equiv 0 \mod \rho^2.$$

Therefore $\rho^k \equiv 0 \mod \rho^{k+1}$ and hence $\rho^k = 0$, contrary to hypothesis. Accordingly $\rho = 0$, so that

$$A_i\phi_i = B_i\phi_i = B_i\psi_i$$

and our theorem is proved.

4. *The canonical form.* The principal idempotent and nilpotent elements assume a very simple form, if the matrix A is taken in the canonical form described below.[9] Let the elementary factors of $A - xE$ be

(33) $$[p_i(x)]^{e_{ij}} \quad (i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, r_i;\ e_{i1} \geqq e_{i2} \geqq \cdots \geqq e_{ir_i}).$$

Further, let P_i denote the companion matrix of $p_i(x)$ and E_{ij} and U_{ij} the

[9] Wedderburn, *op. cit.*, p. 124.

unit matrix and the auxiliary unit matrix of order e_{ij}. Then with the notation of (12) and (13) the matrix

$$N_{ij} = P_i \cdot E_{ij} + E_{\nu_i} \cdot U_{ij} \tag{34}$$

has the single elementary factor $[p_i(x)]^{e_{ij}}$; the diagonal block matrix

$$M_i = [N_{i1}, N_{i2}, \cdots, N_{ir_i}]$$

has the elementary factors $[p_i(x)]^{e_{ij}}$ $(j = 1, 2, \cdots, r_i)$, and the matrix

$$M = [M_1, M_2, \cdots, M_t] = [F_1, M_i, F_2] \tag{35}$$

the elementary factors (33). Accordingly M is similar to A in K and may be taken as the canonical form of A. If A is in the canonical form (35), it follows from Theorem 2 that the matrices ϕ_i, η_i and A_i are given by

$$\begin{aligned} \phi_i &= [0_1, E_{\nu_i} \cdot E_{i1}, E_{\nu_i} \cdot E_{i2}, \cdots, E_{\nu_i} \cdot E_{ir_i}, 0_2], \\ \eta_i &= [0_1, E_{\nu_i} \cdot U_{i1}, E_{\nu_i} \cdot U_{i2}, \cdots, E_{\nu_i} \cdot U_{ir_i}, 0_2], \\ A_i &= [0_1, P_i \cdot E_{i1}, P_i \cdot E_{i2}, \cdots, P_i \cdot E_{ir_i}, 0_2], \end{aligned}$$

where 0_1 and 0_2 are the zero matrices of the same orders respectively as F_1 and F_2 in (35). It is apparent that η_i is nilpotent of index e_{i1} so that $e_{i1} = m_i$.

Corresponding to each particular elementary factor $[p_i(x)]^{e_{ij}}$ is the idempotent element ϕ_{ij}, obtained from ϕ_i by replacing each block except $E_{\nu_i} \cdot E_{ij}$ by the zero matrix, and the nilpotent element η_{ij}, obtained from η_i in the same manner. These idempotent and nilpotent elements reduce to the partial idempotent and nilpotent elements associated with a characteristic number a_i is case $p_1(x)$ is linear.[10]

5. *Applications of the canonical form.* We now proceed to consider the general form of a matrix X commutative with the matrix A, whose elementary factors are given by (33). If $AX = XA$, since ϕ_i is a polynomial in A,

$$\phi_i X = X\phi_i = X_i \qquad (i = 1, 2, \cdots, t), \tag{36}$$

and

$$X = XE = X \sum_{i=1}^{t} \phi_i = \sum_{i=1}^{t} X_i.$$

Further, if

$$A\phi_i X_i = X_i A \phi_i \qquad (i = 1, 2, \cdots, t), \tag{37}$$

$$\begin{aligned} AX &= A \sum_{j=1}^{t} \phi_j \sum_{i=1}^{t} X_i = \sum_{i=1}^{t} A\phi_i X_i, \\ &= \sum_{i=1}^{t} X_i A \phi_i = \sum_{i=1}^{t} X_i \phi_i A = XA. \end{aligned}$$

[10] Cf. Wedderburn, *op. cit.*, p. 42.

Hence, a necessary and sufficient condition, that X be commutative with A, is that (37) be true. If A is in the canonical form (35), it follows from (37) that X is a diagonal block matrix $[X_1, X_2, \cdots, X_t]$, where X_i is a square matrix of the same order as M_i and is commutative with M_i. Since M_i has the single characteristic divisor $p_i(x)$, it will, therefore, be sufficient to consider the case in which A has a single characteristic divisor. We take A in the canonical form

$$A = [N_1, N_2, \cdots, N_r];$$

where N_j is obtained from N_{ij} in (34) by replacing P_i by P, ν_i by ν, e_{ij} by e_j etc. Since $AX = XA$ and A_1 is a polynomial in A, $A_1X = XA_1$. If A_1 and X are considered as matrices with elements, which are matrices of order ν, A_1 is a diagonal matrix, each element being P, and accordingly each element of X is commutative with P and is therefore a polynomial in P. Hence, if θ is a zero of $p(x)$ and $A\{\theta\}$ is the matrix obtained from A by replacing P by θ and E_ν by 1, there is a one to one correspondence between the matrices with elements in K commutative with A and the matrices with elements in $K(\theta)$ commutative with $A\{\theta\}$. The form of a matrix $X\{\theta\}$ commutative with $A\{\theta\}$ is known,[11] and from it we deduce the following result. Let

$$X = (X_{ij}) \qquad (i, j = 1, 2, \cdots, r),$$

where X_i is a matrix with the same number of rows e_i as N_i and the same number of columns e_j as N_j, be a matrix with elements in K commutative with A. Then, if $e_i \geqq e_j = e$,

$$X_{ij} = \begin{pmatrix} G_{ij} \\ 0 \end{pmatrix}, \quad X_{ji} = (0 \cdot G_{ji}), \quad \text{where } G_{ij} \text{ and } G_{ji}$$

are square matrices of order $e\nu$ and are of the form

$$h_1(P) \cdot E_e + h_2(P) \cdot U_e + h_3(P) \cdot U_e{}^2 + \cdots + h_e(P) \cdot U_e{}^{e-1},$$

where $h_i(P)$ is a polynomial in P with coefficients in K.[12]

In conclusion we determine the elementary factors of a matrix polynomial $f(A)$ from the corresponding theorem in the case that K is algebraically closed.[13] It is apparent from the canonical form (35) that the elementary

---

[11] Wedderburn, *op. cit.*, p. 104.

[12] R. C. Trott, *Bulletin of the American Mathematical Society*, January, 1935, abstract no. 95, p. 42.

[13] These results were determined by N. H. McCoy, "On the rational canonical form of a function of a matrix," *American Journal of Mathematics*, vol. 57 (1935), pp. 491-502. The following shows the close connection between his results and the earlier ones.

factors of $f(A)$ are the elementary factors of the matrices $f(N_{ij})$. Hence it will be sufficient to consider the case in which A has the single elementary factor $[p(x)]^m$. We may assume then that A is in the normal form (cf. (13)).

$$A = P \cdot E_m + E_\nu \cdot U_m,$$

so that $A_1 = P \cdot E_m$. Let $f^s(P)$ be the first of the derivatives of $f(P)$ different from zero and let κ be determined by (18), and let

$$l = m - (\kappa - 1)s, \tag{38}$$

so that $1 \leqq l \leqq s$. If θ is a zero of $p(x)$ and $A\{\theta\} = \theta E_m + U_m$, $f^{(s)}(\theta)$ is the first of the derivatives of $f(\theta)$ different from zero, and therefore $f(A\{\theta\})$ has the elementary divisors $[x - f(\theta)]^\kappa$ counted l times and $[x - f(\theta)]^{\kappa-1}$ counted $s - l$ times.[14] Hence there exists a non-singular matrix $C\{\theta\}$ with elements in $K(\theta)$, such that

$$C\{\theta\}A\{\theta\}C^{-1}\{\theta\} = H\{\theta\}, \tag{39}$$

where $H\{\theta\}$ is a diagonal block matrix, consisting of l blocks $f(\theta)E_\kappa + U_\kappa$ and $s - l$ blocks $f(\theta)E_{\kappa-1} + U_{\kappa-1}$. If in (39) θ is replaced by the matrix P,

$$CAC^{-1} = H,^{15}$$

where H consists of l blocks $f(P) \cdot E_\kappa + E_\nu \cdot U_\kappa$ and $s - l$ blocks

$$f(P) \cdot E_{\kappa-1} + E_\nu \cdot U_{\kappa-1}.$$

But, with the notation of Lemma 2, the elementary factors of $f(P) \cdot E_\kappa + E_\nu \cdot U_\kappa$ are $[q(x)]^\kappa$ counted σ times and those of $f(P) \cdot E_{\kappa-1} + E_\nu \cdot U_{\kappa-1}$ are $[q(x)]^{\kappa-1}$ counted σ times. Hence, *if $q(x)$, of degree μ, is the minimal polynomial of $f(P)$, i. e. of $f(A_1)$, the elementary factors of $f(A)$ are $[q(x)]^\kappa$ counted σl times and $[q(x)]^{\kappa-1}$ counted $\sigma(s - l)$ times, where $\sigma = \nu/\mu$.*

Since the integer s in (38) may be defined by the congruences (19), this last result is McCoy's Theorem 1.

THE JOHNS HOPKINS UNIVERSITY.

---

[14] D. E. Rutherford, "On the canonical form of a rational integral function of a matrix," *Proceedings of the Edinburgh Mathematical Society*, ser. 2, vol. 3 (1932), pp. 135-143; C. C. MacDuffee, "On a fundamental theorem in matric theory," *American Journal of Mathematics*, vol. 58 (1936), pp. 504-506.

[15] Since $C\{\theta\}$ is non-singular, C is non-singular. See J. Williamson, "Latent roots of a matrix of special type," *Bulletin of the American Mathematical Society*, vol. 37 (1931), p. 587, Theorem 1.

THE ARITHMETICAL FUNCTION $M(n, f, g)$ AND ITS ASSOCIATES CONNECTED WITH ELLIPTIC POWER SERIES.

By E. T. Bell.

1. *Introduction.* Let $M(n, f, g)$ denote the number of those representations of the integer n as a sum of f squares, precisely g of which are odd and occupy the first g places in the representations, both the arrangement of the squares and the signs of their square roots being relevant in enumerating the representations; also let $N(n, f, g)$ denote the similarly defined function in which the square roots of the odd squares are positive. Then

$$M(n, f, g) = 2^g N(n, f, g). \tag{1.1}$$

In a former paper[1] it was shown that the calculation of $N(n, f, g)$ is an essential step in obtaining the explicit forms of the numerical coefficients occurring in the power series for elliptic functions (in either the Jacobian or the Weierstrassian form). Here we shall give several linear recurrences by which the calculations can be performed expeditiously. Certain sets of these (§§ 9, 10, 11) are complete—no more of the kind given exist. The recurrences introduce functions of divisors, and although these may be evaluated with ease directly, we have also reduced their calculation to linear recurrences. Incidentally some curious results (§ 12) on numbers of representations are noted.

The initial values must be noticed. Refer to (2.1) for ϵ.

(1.2) $M(s, 1, 0) = 0$ if $s \not\equiv 0 \bmod 4$,
$M(s, 1, 0) = 2\epsilon(s/4)$ if $s \equiv 0 \bmod 4$;

(1.3) $M(s, g, g) = 0$ if $s \not\equiv g \bmod 8$,
$M(8s + 1, 1, 1) = 2\epsilon(8s + 1)$;

(1.4) $M(s, f, g) = 0$ if $g > f$; in all of which $s \geqq 0$;

(1.5) $M(0, f, 0) = 1$, $f \geqq 0$; $M(n, 0, 0) = 0$, $n > 0$.

If $R_r(n)$ denotes the total number of representations of n as a sum of r squares,

[1] *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 841-852. The functions M, N are also useful in the theory of numbers; see R. D. James, *American Journal of Mathematics*, vol. 58 (1936), pp. 536-544, where further references are given.

(1.6) $M(4n, r, 0) = R_r(n)$.

2. *Arithmetical functions.* For convenient reference we collect here the definitions of the arithmetical functions occurring in the sequel.

In the following definitions n, m, α, r, a, b denote integers; $n > 0$; $\alpha \geqq 0$; $m > 0$ is odd; $n = 2^\alpha m$; $r, a, b \geqq 0$.

(2.1) $\epsilon(n) = 1$ or 0 according as n is or is not a square; $\epsilon(0) = \frac{1}{2}$.

(2.2) $(-1|m) \equiv (-1)^{(m-1)/2}$.

(2.3) $\zeta_r(n) =$ the sum of the r-th powers of all the divisors of n.

(2.4) $\zeta'_r(n) \equiv \zeta_r(m)$.

(2.5) $\xi_r(n) =$ the excess of the sum of the r-th powers of all those divisors of n that are of the form $4k+1$ over the like sum for the divisors of the form $4k+3$; $\xi(n) \equiv \xi_0(n)$.

(2.6) $\xi'_r(n) =$ the excess of the sum of the r-th powers of all those divisors of n whose conjugates are of the form $4k+1$ over the like sum in which the conjugates are of the form $4k+3$. (If $n = d\delta$, d, δ integers > 0, d, δ are called conjugate divisors of n.)

(2.7) $\xi''_r(n) \equiv \xi_r(n) - \xi'_r(n)$.

(2.8) $\alpha_r(n) \equiv n^r\zeta'_{-r}(n)$, $=$ the sum of the r-th powers of all those divisors of n whose conjugates are odd.

(2.9) $\lambda_r(n) \equiv [1 + 2(-1)^n]\zeta'_r(n)$.

(2.10) $\beta_r(n) \equiv 4\xi'_r(n) - \xi_r(n)$.

(2.11) $\zeta''_r(n) =$ the sum of the r-th powers of all the even divisors of n.

(2.12) $\rho_r(n) \equiv \zeta'_r(n) - \zeta''_r(n)$.

(2.13) $\omega(n) \equiv 2^\alpha\zeta_1(m)$.

(2.14) $\theta_{a,b}(n)$
$$\equiv [3a + b + (-1)^n(3a - b)]\zeta'_1(n) - (a+b)[1 + (-1)^n]\omega(n).$$

(2.15) $\gamma(n) \equiv 2\zeta'_1(n) - \zeta_1(n) = (3 - 2^{\alpha+1})\zeta_1(m)$.

As immediate consequences of the definitions we have

(2.16)
$$\xi_r(n) = \xi_r(m),\ \xi'_r(m) = (-1|m)\xi_r(m);$$
$$\xi'_r(n) = (-1|m)2^{\alpha r}\xi_r(m),\ \xi''_r(n) = [1 - (-1|m)2^{\alpha r}]\xi_r(m);$$
$$\beta_r(n) = [2^{\alpha r+2}(-1|m) - 1]\xi_r(m);$$
$$\xi(4n-1) = 0;\ \omega(2n) = 2\omega(n).$$

(2.17)
$$\theta_{a,b}(2n) = 6a\zeta'_1(n) - 4(a+b)\omega(n),$$
$$\theta_{a,b}(m) = 2b\zeta_1(m).$$

3. *Connection of M, N with elliptic power series.* By (1.1) it is suffi-

cient to show the connection for either M or N. For comparison with the paper cited [2] we choose N. As before, m is odd.

$$(3.1)\qquad \operatorname{cn} x = 1 + \sum_{s=1}^{\infty} (-1)^s Q_s(k^2)[x^{2s}/(2s)!],\quad Q_s(k^2) \equiv \sum_{r=0}^{s-1} q_r(s) k^{2r};$$

$$\sum_{r=0}^{(m-1)/2} 2^{4r} N(2m, 4s+2, 4r+2) q_r(s) = \xi_{2s}(m).$$

$$(3.2)\qquad \operatorname{sn} x = \sum_{s=0}^{\infty} \frac{(-1)^s P_s(k^2)}{(2s+1)!} x^{2s+1},\quad P_s(k^2) \equiv \sum_{r=0}^{s} p_r(s) k^{2r};$$

$$\sum_{r=0}^{s} 2^{4r} p_r(s) N(2m, 4s+4, 4r+2) = \zeta_{2s+1}(m).$$

$$(3.3)\qquad \operatorname{dn} x = 1 + \sum_{s=1}^{\infty} (-1)^s R^s(k^2)[x^{2s}/(2s)!],\quad R_s(k^2) \equiv \sum_{j=0}^{s-1} r_j(s) k^{2j+2};$$

$$\sum_{j=0}^{n-1} 2^{4j} N(4n, 4s+2, 4j+4) r_j(s) = 2^{2s-2}\xi_{2s}(n).$$

There are similar results for $x/\operatorname{sn} x$, $x^2/\operatorname{sn}^2 x$, $\wp(x)$, or any elliptic function or quotient of theta functions. From the recurrences for the $q_r(s)$, $p_r(s)$, $r_j(s)$ the explicit forms of the $Q_s(k^2)$, $P_s(k^2)$, $R_s(k^2)$ are calculated in terms of functions N, ξ_{2s}, ζ_{2s+1}. Thus (see the paper cited)

$$q_0(s) = 1,\ 2^4 q_1(s) = 3^{2s} - 8s - 1,$$
$$2^8 q_2(s) = 5^{2s} - 8(s-1)3^{2s} + 32s^2 - 48s - 9,$$

and so on. The general form of $q_r(s)$ (also of the other coefficients) is readily inferred (*loc. cit.*) from the recurrences involving N. For the first few coefficients $q_r(s)$, etc., the necessary N's are easily calculated directly; for a systematic evaluation this more or less tentative method is impracticable.

It may be noted that an interesting sequence of successive approximations to the values of elliptic functions is obtained by retaining only powers of k^2 not exceeding the t-th, for $t = 0, 1, 2, \cdots$. The series so truncated at the t-th power of k^2 can be summed in finite form.

4. *Generating functions.* In the usual notation for the elliptic theta functions we have (4.1)–(4.3), in which sums and products refer to all integers $n > 0$, to all integers $\nu \gtreqless 0$, to all odd integers $m > 0$, or to all odd integers $\mu \gtrless 0$ respectively.

$$(4.1)\qquad Q_r \equiv Q_r(q):$$
$$Q_0 = \Pi(1 - q^{2n}),\qquad Q_1 = \Pi(1 + q^{2n}),$$
$$Q_2 = \Pi(1 + q^m),\qquad Q_3 = \Pi(1 - q^m).$$

[2] The following misprints occur: p. 842 (1), for x^2 read x^{2s}; p. 843, last line, for 2^{2s} read 3^{2s}.

$$
\begin{aligned}
(4.2)\qquad & \vartheta_r \equiv \vartheta_r(q): \\
& \vartheta_0 = 1 + 2\Sigma(-1)^n q^{n^2} = \Sigma(-1)^\nu q^{\nu^2}; \\
& \vartheta'_1 = 2\Sigma(-1|m)\, m q^{m^2/4} = \Sigma(-1|\mu)\,\mu q^{\mu^2/4}; \\
& \vartheta_2 = 2\Sigma q^{m^2/4} = \Sigma q^{\mu^2/4}; \\
& \vartheta_3 = 1 + 2\Sigma q^{n^2} = \Sigma q^{\nu^2}.
\end{aligned}
$$

$$
\begin{aligned}
(4.3)\qquad & Q_1Q_2Q_3 = 1, & \vartheta'_1 = \vartheta_0\vartheta_2\vartheta_3; \\
& \vartheta_0 = Q_0Q_3^2, & \vartheta'_1 = 2q^{1/4}Q_0^3, \\
& \vartheta_2 = 2q^{1/4}Q_0Q_1^2, & \vartheta_3 = Q_0Q_2^2.
\end{aligned}
$$

The sum of a odd squares is of the form $8s + a$; the sum of b even squares is of the form $4s$. Thus, from the definition (§ 1) of M,

$$
(4.4)\qquad \vartheta_2^a(q^4)\vartheta_3^b(q^4) = \sum_{s=0}^{\infty} q^{4s+a} M(4s + a, a + b, a), \quad (a, b \text{ integers} \geqq 0).
$$

Hence, from (4.3),

$$
(4.5)\qquad 2^a Q_0^{a+b} Q_1^{2a} Q_2^{2b} = \sum_{s=0}^{\infty} q^s M(4s + a, a + b, a).
$$

Define the operation Δ by

$$
(4.6)\qquad \Delta F(q) = q(d/dq) \log F(q).
$$

Then $\Delta[F(q)\cdots G(q)] = \Delta F(q) + \cdots + \Delta G(q)$. A short calculation gives (4.7)–(4.9), in which Σ refers to all integers $n > 0$, (see the definitions in § 2).

$$
\begin{aligned}
(4.7)\qquad & \Delta Q_0 = -2\Sigma q^{2n}\zeta_1(n), & \Delta Q_1 = 2\Sigma q^{2n}\zeta'_1(n), \\
& \Delta Q_2 = -\Sigma q^n(-1)^n\zeta'_1(n), & \Delta Q_3 = -\Sigma q^n\zeta'_1(n);
\end{aligned}
$$

hence, by (4.6), (4.3),

$$
(4.8)\qquad \Delta(Q_0^{a+b}Q_1^{2a}Q_2^{2b}) = \Sigma q^n \theta_{a,b}(n);
$$

$$
\begin{aligned}
(4.9)\qquad & \Delta\vartheta_0 = -2\Sigma q^n\omega(n), & 4\Delta\vartheta'_1 = 1 - 24\Sigma q^{2n}\zeta_1(n), \\
& 4\Delta\vartheta_2 = 1 + 8\Sigma q^{2n}\gamma(n), & \Delta\vartheta_3 = -2\Sigma q^n(-1)^n\omega(n).
\end{aligned}
$$

The next are equivalent to the classical theorems on numbers of representations as sums of 2, 4, 6, 8 squares.[3] By (4.4) the two statements in each of (4.10)–(4.23) express the same theorem. The summations refer to all integers $n > 0$ or to all odd integers $m > 0$, and s is an integer $\geqq 0$.

[3] All are collected in my paper, *American Journal of Mathematics*, vol. 42 (1920), pp. 168-188. They follow at once from theorems of Jacobi in the *Fundamenta Nova*.

(4.10) $\vartheta_2^2 = 4\Sigma q^{m/2}\xi(m), \quad M(8s+2, 2, 2) = 4\xi(4s+1).$
(4.11) $\vartheta_2^4 = 16\Sigma q^m \zeta_1(m), \quad M(8s+4, 4, 4) = 16\zeta_1(2s+1).$
(4.12) $\vartheta_2^6 = -4\Sigma q^{m/2}\xi_2''(m), \quad M(8s+6, 6, 6) = -4\xi_2''(4s+3).$
(4.13) $\vartheta_2^8 = 256\Sigma q^{2n}\alpha_3(n), \quad M(8s+8, 8, 8) = 256\alpha_3(4s+4).$
(4.14) $\vartheta_3^2 = 1 + 4\Sigma q^n \xi(n), \quad M(4n, 2, 0) = 4\xi(n).$
(4.15) $\vartheta_3^4 = 1 + 8\Sigma q^n(-1)^n\lambda_1(n), \quad M(4n, 4, 0) = 8(-1)^n\lambda_1(n).$
(4.16) $\vartheta_3^6 = 1 + 4\Sigma q^n \beta_2(n), \quad M(4n, 6, 0) = 4\beta_2(n).$
(4.17) $\vartheta_3^8 = 1 - 16\Sigma q^n(-1)^n\rho_3(n), \quad M(4n, 8, 0) = -16(-1)^n\rho_3(n).$
(4.18) $\vartheta_2\vartheta_3 = 2\Sigma q^{m/4}\xi(m), \quad M(4s+1, 2, 1) = 2\xi(4s+1).$
(4.19) $\vartheta_2^2\vartheta_3^2 = 4\Sigma q^{m/2}\zeta_1(m), \quad M(4s+2, 4, 2) = 4\zeta_1(2s+1).$
(4.20) $\vartheta_2^3\vartheta_3^3 = -\tfrac{1}{2}\Sigma q^{m/4}\xi_2''(m), \quad M(4s+3, 6, 3) = -\tfrac{1}{2}\xi_2''(4s+3).$
(4.21) $\vartheta_2^4\vartheta_3^4 = 16\Sigma q^n\alpha_3(n), \quad M(4s+4, 8, 4) = 16\alpha_3(s+1).$
(4.22) $\vartheta_2^2\vartheta_3^4 = 4\Sigma q^{m/2}\xi'_2(m), \quad M(4s+2, 6, 2) = 4\xi'_2(2s+1).$
(4.23) $\vartheta_2^4\vartheta_3^2 = 16\Sigma q^n\xi'_2(n), \quad M(4s+4, 6, 4) = 16\xi'_2(s+1).$

5. *Recurrence for $M(n, f, g)$ with f, g constant throughout.* Operating with Δ on both sides of (4.5), using (4.8), and comparing coefficients of like powers of q in the result, we get

(5.1) $$nM(4n+a, a+b, a) = \sum_{s=1}^{n} \theta_{a,b}(s)M(4n+a-4s, a+b, a).$$

With the initial value $M(a, a+b, a) = 2^a$, (5.1) suffices to calculate all $M(n, f, g)$ with f, g constant throughout; necessarily $n \equiv g \bmod 4$:

(5.2) $$nM(4n+g, f, g) = \sum_{s=1}^{n} \theta_{g,f-g}(s)M(4n+g-4s, f, g).$$

From (2.14), (2.17) we have

(5.3) $$\theta_{g,f-g}(n) = [2g+f+(-1)^n(4g-f)]\zeta'_1(n) - f[1+(-1)^n]\omega(n),$$
$$\theta_{g,f-g}(2n) = 6g\zeta'_1(n) - 4f\omega(n), \quad \theta_{g,f-g}(m) = 2(f-g)\zeta_1(m).$$

Recurrences for the associated functions ζ'_1 (by means of λ_1) and ω are given in §§ 7, 8. This type is continued in § 10.

6. *Arguments n in $N(n, f, g)$ decreasing by squares.* Multiply the left of (4.4) by the second form of $\vartheta_2(q^4)$ in (4.2), and the right by the first form; do the like with $\vartheta_3(q^4)$. Then, for $n \geqq 0$,

(6.1) $$M(4n+a+1, a+b+1, a+1)$$
$$= 2\Sigma M(4n+a+1-m^2, a+b, a),$$

(6.2) $$M(4n+a, a+b+1, a)$$
$$= M(4n+a, a+b, a) + 2\Sigma M(4n+a-4s^2, a+b, a),$$

where the summations refer to $m = 1, 3, 5, \cdots$, and $s = 1, 2, 3, \cdots$ respectively, continuing so long as the arguments are $\geqq 0$. Hence

$$(6.3)\quad M(4n+g,f,g) = 2\Sigma M(4n+g-m^2, f-1, g-1),$$

$$(6.4)\quad M(4n+g,f,g) = M(4n+g,f-1,g) + 2\Sigma M(4n+g-4s^2, f-1, g).$$

The sum of u odd squares is of the form $8t+u$ $(t \geqq 0)$; the sum of v even squares is of the form $4t$. Hence

$$(6.41)\quad \vartheta_2^u(q^4) = \sum_{t=0}^{\infty} q^{8t+u} M(8t+u, u, u), \quad \vartheta_3^v = \sum_{t=0}^{\infty} q^{4t} M(4t, v, 0).$$

Operating on these with Δ and proceeding as in § 5 we get

$$(6.5)\quad \Sigma[8n - (u+1)(m^2-1)] M(8n+u+1-m^2, u, u) = 0,$$

$$(6.6)\quad nM(4n, v, 0) + 2\Sigma[n - (v+1)s^2] M(4n-4s^2, v, 0) = 2v\epsilon(n)n,$$

the summations with respect to m, s being as in (6.3), (6.4), with $s^2 < n$ in (6.6). We shall generalize (6.5), (6.6) in § 10.

7. *Recurrences for the functions in* § 2. Applying (4.10)–(4.13) to (6.5) we find the following, in which $n > 0$, and the summations refer to $m = 1, 3, 5, \cdots$.

$$(7.1)\quad \Sigma[8n - 3(m^2-1)]\xi\ (4n+1-\tfrac{1}{2}(m^2-1)) = 0,$$
$$(7.2)\quad \Sigma[8n - 5(m^2-1)]\zeta_1\ (2n+1-\tfrac{1}{4}(m^2-1)) = 0,$$
$$(7.3)\quad \Sigma[8n - 7(m^2-1)]\xi_2''(4n+3-\tfrac{1}{2}(m^2-1)) = 0,$$
$$(7.4)\quad \Sigma[8n - 9(m^2-1)]\alpha_3\ (4n+4-\tfrac{1}{2}(m^2-1)) = 0.$$

In the same way, from (4.14)–(4.17) and (6.6) we get the following recurrences with $n > 0$, and Σ referring to $t = 1, 2, 3, \cdots$, with $t^2 < n$.

$$(7.5)\quad n\xi(n) + 2\Sigma(n-3t^2)\xi(n-t^2) = \epsilon(n)n,$$
$$(7.6)\quad n\lambda_1(n) + 2\Sigma(-1)^t(n-5t^2)\lambda_1(n-t^2) = (-1)^n\epsilon(n)n,$$
$$(7.7)\quad n\beta_2(n) + 2\Sigma(n-7t^2)\beta_2(n-t^2) = 3\epsilon(n)n,$$
$$(7.8)\quad n\rho_3(n) + 2\Sigma(-1)^t(n-9t^2)\rho_3(n-t^2) = (-1)^{n-1}\epsilon(n)n.$$

8. *Continuation of* § 7. Operating with Δ on the first forms of the thetas in (4.2) and comparing with (4.9) we find

$$(8.1)\quad \omega(n) + 2\Sigma(-1)^t\omega(n-t^2) = \epsilon(n)(-1)^{n-1}n;$$

$$(8.2)\quad \zeta_1(n) + \Sigma(-1)^t(2t+1)\zeta_1(n-\tfrac{1}{2}t(t+1))$$
$$= \begin{cases} 0 \text{ if } n \neq \tfrac{1}{2}h(h+1) \quad (h>0), \\ -\tfrac{1}{6}(-1)^h h(h+1)(2h+1) \text{ if } n = \tfrac{1}{2}h(h+1); \end{cases}$$

$$(8.3)\qquad \gamma(n)+\Sigma\gamma(n-\tfrac{1}{2}t(t+1))=\begin{cases}0, & n\neq\tfrac{1}{2}h(h+1),\\ n, & n=\tfrac{1}{2}h(h+1).\end{cases}$$

The summations refer to $t=1,2,3,\cdots$; and $n>0$. This is continued in § 11.

9. *Recurrences with arguments n of $M(n,f,g)$ in arithmetic progression.* Applying (4.4), (6.41) to $\vartheta_2^{g+u}\vartheta_3^b=\vartheta_2^u\times\vartheta_2^g\vartheta_3^b$ we get

$$M(4n+g+u,f+u,g+u)=\Sigma M(8s+u,u,u)M(4n+g-8s,f,g),$$

with $n\geqq 0$ and Σ referring to $s=0,1,\cdots,[n/2]$.

Hence, by (4.10)–(4.13), for the same n, s,

$$(9.1)\quad M(4n+g+2,f+2,g+2)=4\Sigma\xi\,(4s+1)M(4n+g-8s,f,g);$$
$$(9.2)\quad M(4n+g+4,f+4,g+4)=16\Sigma\zeta_1\,(2s+1)M(4n+g-8s,f,g);$$
$$(9.3)\quad M(4n+g+6,f+6,g+6)=-4\Sigma\xi_2''(4s+3)M(4n+g-8s,f,g);$$
$$(9.4)\quad M(4n+g+8,f+8,g+8)=256\Sigma\alpha_3\,(4s+4)M(4n+g-8s,f,g).$$

Similarly, $\vartheta_2^a\vartheta_3^{b+v}=\vartheta_2^a\vartheta_3^b\times\vartheta_3^v$;

$$M(4n+g,f+2,g)=M(4n+g,f,g)+4\Sigma\xi(s)M(4n+g-4s,f,g),$$

for $n\geqq 0$, $s=1,\cdots,n$, *with the convention that a sum in which the lower limit exceeds the upper is vacuous.* Hence, from (4.14)–(4.17), for the same n, s,

$$(9.5)\quad M(4n+g,f+2,g)=M(4n+g,f,g)+4\Sigma\xi(s)M(4n+g-4s,f,g);$$
$$(9.6)\quad M(4n+g,f+4,g)=M(4n+g,f,g)+8\Sigma(-1)^s\lambda_1(s)M(4n+g-4s,f,g);$$
$$(9.7)\quad M(4n+g,f+6,g)=M(4n+g,f,g)+4\Sigma\beta_2(s)M(4n+g-4s,f,g);$$
$$(9.8)\quad M(4n+g,f+8,g)=M(4n+g,f,g)-16\Sigma(-1)^s\rho_3(s)M(4n+g-4s,f,g).$$

From $\vartheta_2^{a+c}\vartheta_3^{b+c}=\vartheta_2^a\vartheta_3^b\times\vartheta_2^c\vartheta_3^c$,

$$M(4n+g+c,f+2c,g+c)=\Sigma M(4s+c,2c,c)M(4n+g-4s,f,g),$$

for $n\geqq 0$, $s=0,\cdots,n$. Hence for the same n, s, (4.18)–(4.21) give

$$(9.9)\quad M(4n+g+1,f+2,g+1)=2\Sigma\xi\,(4s+1)M(4n+g-4s,f,g);$$
$$(9.10)\quad M(4n+g+2,f+4,g+2)=4\Sigma\zeta_1\,(2s+1)M(4n+g-4s,f,g);$$
$$(9.11)\quad 2M(4n+g+3,f+6,g+3)=-\Sigma\xi_2''(4s+3)M(4n+g-4s,f,g);$$
$$(9.12)\quad M(4n+g+4,f+8,g+4)=16\Sigma\alpha_3\,(s+1)M(4n+g-4s,f,g).$$

Finally, $\vartheta_2^{a+u}\vartheta_3^{b+v} = \vartheta_2^a\vartheta_3^b \times \vartheta_2^u\vartheta_3^v$;

$$M(4n+g+u, f+u+v, g+u) = \Sigma M(4s+u, u+v, u) M(4n+g-4s, f, g),$$

for $n \geqq 0$, and $s = 0, \cdots, n$. For the same n, s, we get from (4. 22), (4. 23),

$$(9.13) \quad M(4n+g+2, f+6, g+2) = 4\Sigma\xi'_2(2s+1) M(4n+g-4s, f, g);$$
$$(9.14) \quad M(4n+g+4, f+6, g+4) = 16\Sigma\xi'_2(s+1) M(4n+g-4s, f, g).$$

10. *Continuation of* § 5. By (4. 4),

$$\sum_{s=0}^{\infty} q^{4s+gr} M(4s+gr, fr, gr) = [\sum_{t=0}^{\infty} q^{4t+g} M(4t+g, f, g)]^r$$

is equivalent to the identity

$$\vartheta_2^{gr}\vartheta_3^{(f-g)r} = (\vartheta_2^g\vartheta_3^{f-g})^r;$$

hence, operating throughout the first with Δ, we find

$$(10.1) \quad \Sigma[n-(r+1)s] M(4s+g, f, g) M(4n+gr-4s, fr, gr) = 0,$$

valid for $n \geqq 0$, the sum referring to $s = 0, \cdots, n$. This contains as special cases (6. 5), (6. 6). To obtain (6. 5) from (10. 1) take $r = u$, $f = g = 1$, and note that s, n must then be even; to obtain (6. 6) take $r = v$, $f = 1$, $g = 0$, and apply (1. 2). In (10. 1) take $f = g$, and apply (1. 3),

$$(10.2) \quad \Sigma[n-(r+1)s] M(8s+g, g, g) M(8n+gr-8s, gr, gr) = 0.$$

For $g = 0$, (10. 1) becomes, by (1. 5),

$$(10.3) \quad nM(4n, fr, 0) + 4\Sigma[n-(r+1)s] M(4s, f, 0) M(4n-4s, fr, 0) = rnM(4n, f, 0),$$

in which $n \geqq 0$ and Σ refers to $s = 1, \cdots, n-1$.

From (10. 2) and (4. 10)–(4. 13) we have the following, n and Σ being as in (10. 1),

$$(10.4) \quad \Sigma[n-(r+1)s]\xi \quad (4s+1) M(8n+2r-8s, 2r, 2r) = 0,$$
$$(10.5) \quad \Sigma[n-(r+1)s]\zeta_1 \ (2s+1) M(8n+4r-8s, 4r, 4r) = 0,$$
$$(10.6) \quad \Sigma[n-(r+1)s]\xi_2''(4s+3) M(8n+6r-8s, 6r, 6r) = 0,$$
$$(10.7) \quad \Sigma[n-(r+1)s]\alpha_3 \ (4s+4) M(8n+8r-8s, 8r, 8r) = 0.$$

Taking $f = 2, 4, 6, 8$ in (10. 3) and referring to (4. 14)–(4. 17), we find (10. 8)–(10. 11), in which n, Σ are as in (10. 3),

(10.8) $nM(4n, 2r, 0) + 4\Sigma[n - (r+1)s]\xi(s)M(4n - 4s, 2r, 0) = 4rn\xi(n)$,

(10.9) $nM(4n, 4r, 0) + 8\Sigma(-1)^s[n - (r+1)s]\lambda_1(s)M(4n - 4s, 4r, 0)$
$= 8(-1)^n rn\lambda_1(n)$,

(10.10) $nM(4n, 6r, 0) + 4\Sigma[n - (r+1)s]\beta_2(s)M(4n - 4s, 6r, 0)$
$= 4rn\beta_2(n)$,

(10.11) $nM(4n, 8r, 0) - 16\Sigma(-1)^s[n - (r+1)s]\rho_3(s)M(4n - 4s, 8r, 0)$
$= -16(-1)^n rn\rho_3(n)$.

From (10.1) and (4.18)–(4.23),

(10.12) $\Sigma[n - (r+1)s]\xi\ (4s+1)M(4n + r - 4s, 2r, r) = 0$,
(10.13) $\Sigma[n - (r+1)s]\zeta_1\ (2s+1)M(4n + 2r - 4s, 4r, 2r) = 0$,
(10.14) $\Sigma[n - (r+1)s]\xi_2''(4s+3)M(4n + 3r - 4s, 6r, 3r) = 0$,
(10.15) $\Sigma[n - (r+1)s]\alpha_3\ (s+1)M(4n + 4r - 4s, 8r, 4r) = 0$,
(10.16) $\Sigma[n - (r+1)s]\xi'_2\ (2s+1)M(4n + 2r - 4s, 6r, 2r) = 0$,
(10.17) $\Sigma[n - (r+1)s]\xi'_2\ (s+1)M(4n + 4r - 4s, 6r, 4r) = 0$,

with Σ, n as in (10.1).

11. *Continuation of* § 8. In (10.4) take $r = 2, 3, 4$; in (10.5), $r = 2$; in (10.6), $r = 1$, and apply (4.10)–(4.13). Then, for $n \geqq 0$ and the summation referring to $s = 0, 1, \cdots, n$, we have (11.1)–(11.5). The obvious omitted possibilities give relations which are trivially true, and likewise for subsequent omissions.

(11.1) $\Sigma(n - 3s)\xi\ (4s+1)\zeta_1\ (2n + 1 - 2s) = 0$,
(11.2) $\Sigma(n - 4s)\xi\ (4s+1)\xi_2''(4n + 3 - 4s) = 0$.
(11.3) $\Sigma(n - 5s)\xi\ (4s+1)\alpha_3\ (4n + 4 - 4s) = 0$,
(11.4) $\Sigma(n - 3s)\zeta_1\ (2s+1)\alpha_3\ (4n + 4 - 4s) = 0$,
(11.5) $\Sigma(n - 7s)\xi_2''(4s+3)\xi_2''(4n + 3 - 4s) = 0$.

In (10.8) take $r = 2, 3, 4$; in (10.9) take $r = 2$, and apply (4.14)–(4.17). Then, for n, Σ as in (10.3), we have

(11.6) $n\lambda_1(n) + 4\Sigma(-1)^s(n - 3s)\xi(s)\lambda_1(n - s) = (-1)^n\xi(n)$,
(11.7) $n\beta_2(n) + 4\Sigma(n - 4s)\xi(s)\beta_2(n - s) = 3n\xi(n)$,
(11.8) $n\rho_3(n) + 4\Sigma(-1)^s(n - 5s)\xi(s)\rho_3(n - s) = -(-1)^n n\xi(n)$,
(11.9) $n\rho_3(n) + 8\Sigma(n - 3s)\lambda_1(s)\rho_3(n - s) = -n\lambda_1(n)$.

The formulas (10.12)–(10.17) furnish no new results of the above kind.

12. *Polynomial forms of M.* Taking $n = 0, 1, 2, \cdots$ in (5.1) and referring to (2.17) we find

$$
\begin{aligned}
(12.1) \qquad M(a, a+b, a) &= 2^a, \\
M(4+a, a+b, a) &= 2^{a+1}b, \\
M(8+a, a+b, a) &= 2^a(2b^2 - 2b + a), \\
3M(12+a, a+b, d) &= 2^{a+1}b(2b^2 - 6b + 3a + 4), \\
6M(16+a, a+b, a) & \\
&= 2[4b^4 - 24b^3 + 4(3a+11)b^2 - 12(a+1)b + 3a(a-1)],
\end{aligned}
$$

the general nature of $M(4n + a, a + b, a)$ (proved by mathematical induction) being evident. Take $a = 0$ and refer to (1.6); then

(12.2) *The number of representations of n as a sum of b squares is expressible as a polynomial of degree n in b with rational coefficients.*

Taking $b = 0$ we get

(12.3) *The number of representations of $8n + a$ as a sum of a odd squares is expressible as 2^a times a polynomial of degree n in a with rational coefficients.*

California Institute of Technology,
Pasadena, California.

REPRESENTATION AS SUMS OF MULTIPLES OF GENERALIZED POLYGONAL NUMBERS.

By L. W. Griffiths.

1. *Introduction.* The problem is a generalization of that of representation as sums of multiples of squares,[1] as sums of polygonal numbers,[2] of extended polygonal numbers,[3] or of generalized polygonal numbers.[4]

In my first two papers [5] there are summaries of all of these facts on representation, and illustrations.

The generalized polygonal numbers of order $m+2$ are the values of

$$g(x) = x + m(x^2 - x)/2 \quad \text{for} \quad x = 0, \pm 1, \pm 2, \cdots \tag{1}$$

with m a fixed positive integer. The following notations are used: n and $a_1, \cdots, a_n$ are positive integers, $g_i = g(x_i)$, and

$$\begin{aligned} f &= a_1 g_1 + \cdots + a_n g_n = (a_1, \cdots, a_n), \quad 1 \leqq a_1 \leqq \cdots \leqq a_n, \\ f_4 &= (a_5, \cdots, a_n), \\ w_k &= a_1 + \cdots + a_k \; (1 \leqq k \leqq n), \quad w = w_n. \end{aligned} \tag{2}$$

The positive integer N is represented by the function f when there are integers $x_1, \cdots, x_n$ such that $f = N$. In section 4 of II, conditions were found that f represent the integers $0, 1, \cdots, 34m - 16$. In section 5 of II, for certain functions f satisfying these necessary conditions, there was found a positive integer M, depending only on m and f, such that f represents every integer $N \geqq M$. For other functions satisfying the necessary conditions the methods there employed gave no conclusion. For the remaining functions satisfying the necessary conditions there was no generalized Cauchy lemma on simultaneous representation, and hence the methods there employed were inapplicable.

In this paper use is made of the new facts on simultaneous representation due to Dickson [6] and, for the case $(1, 1, 2, 4)$ with b divisible by four,

---

[1] Dickson, *Bulletin of the American Mathematical Society*, vol. 33 (1927), pp. 63-70.

[2] Cauchy, *Œuvres*, ser. 2, vol. 6, pp. 320-353.

[3] Dickson, *American Journal of Mathematics*, vol. 50 (1928), pp. 1-48.

[4] Dickson, *Journal de Mathématiques*, ser. 9, vol. 7 (1928), Theorems 11-15.

[5] *Annals of Mathematics*, ser. 2, vol. 31 (1930), pp. 1-12, and *American Journal of Mathematics*, vol. 55 (1933), pp. 102-110. These papers will be cited as I and II.

[6] *American Journal of Mathematics*, vol. 56 (1934), pp. 512-528.

here first treated in section 7. There are five fundamental cases, and the details of proof are usually omitted in the last four cases. Each case is summarized in a theorem: *if f satisfies the necessary conditions mentioned, then there is exhibited a positive integer M, depending only on m and f, such that f represents every integer $N \geqq M$.*

2. *The functions* $(1, 1, 2, 3, \cdots)$. By (1) and (2) the positive integer N is represented by f if there exist positive integers a, b, r such that $N = r + b + m(a - b)/2$, that r is represented by f_4, and that the equations $a = x^2 + y^2 + 2z^2 + 3w^2$, $b = x + y + 2z + 3w$ have a solution in integers $\geqq 0$. Sufficient conditions for the existence of these integers x, y, z, w are [7]

$$(3) \qquad a \equiv b \pmod 2, \quad b^2 \leqq ta, \quad (t-1)a < b^2 + 2b + t, \quad b \geqq 0,$$

$$(4) \qquad 6(7a - b^2) \neq 49^s(7n + e), \quad e = 3, 5, 6.$$

In this paper there is found an integer M, depending only on m and f, such that if N is an integer $\geqq M$ then there exist integers a, b, r satisfying (3) and (4), and such that $N = r + b + m(a - b)/2$ and f_4 represents r. Hence to prove the universality of the functions f, satisfying the necessary conditions mentioned in the introduction, it remains to prove that f represents integers $34m - 16 < N < M$.

Dickson proved that (4) is satisfied if either a or b is prime to $t = 7$. The first step in my new method is to obtain conditions which are equivalent to (4) when $a \equiv b \equiv 0 \pmod 7$. Write $b = 7^iB$ with $B \not\equiv 0 \pmod 7$, and $a = 7^{2h}A$ or $7^{2h-1}A$ with $A \not\equiv 0 \pmod 7$. Then it is easily seen that $6(7a - b^2) = 49^s(7n + e)$ with $e = 3, 5, 6$ if and only if one of (5) is satisfied:

$$(5) \qquad \begin{array}{lll} a = 7^{2h-1}A, & 1 \leqq h < i, & A \equiv -e \pmod 7); \quad \text{or,} \\ a = 7^{2h-1}A, & 1 \leqq h = i, & A \equiv B^2 - e \pmod 7). \end{array}$$

That is, (4) is satisfied if and only if one of (6) or (7) is satisfied:

$$(6) \qquad \begin{array}{lll} a = 7^{2h-1}A, & 1 \leqq h < i, & A \equiv e \pmod 7); \quad \text{or} \\ a = 7^{2h-1}A, & 1 \leqq h = i, & A \not\equiv B^2 - e \pmod 7); \quad \text{or} \\ a = 7^{2h-1}A, & 1 \leqq i < h; & \end{array}$$

$$(7) \qquad a = 7^{2h}A.$$

The next step leading to the desired integers a, b, r is a consideration

[7] Dickson, *American Journal of Mathematics*, vol. 56, loc. cit. Here, by definition, $t = a_1 + a_2 + a_3 + a_4$.

of the integers, between 0 and $m-2-t$, not represented by f_4. Among the necessary conditions which these functions f satisfy are $w=m-2$ and $a_k \leqq w_{k-1}+1$ $(5 \leqq k \leqq n)$. Hence, defining $w'_j = a_5 + \cdots + a_{j+4}$, $w'_j = w_{j-4} - t$ $(1 \leqq j \leqq n-4)$ and $a_{j+4} \leqq w'_{j-1} + t + 1$. Hence, by applying Lemma 1 of my first paper I (valid, as proved in II, page 107), we have

LEMMA 1. *If η is an integer between 0 and $m-2-t$ then f_4 represents at least one of the integers $\eta, \eta-1, \cdots, \eta-t$.*

Next I shall prove that if $N \geqq 28$ then there exist integers a, b, r satisfying all the conditions except perhaps (3_2) and (3_3).

LEMMA 2. *Let f satisfy the necessary conditions mentioned above. Let ξ be an integer not divisible by t, and $4t \leqq \xi$. Let $N \geqq \xi$. Then there are integers a, b, r, each $\geqq 0$, such that $N = r + b + m(a-b)/2$, $a \equiv b \pmod 2$, (4) is satisfied, and $f_4 = r \leqq m-2-t$.*

For there are integers g and ρ such that $N = mg + \xi + \rho$, $g \geqq 0$, $0 \leqq \rho \leqq m-1$. If $f_4 \neq \rho < m-2-t$, then by Lemma 1 we can write $N = mg + \xi + i + (\rho - i)$ where i is one of $0, 1, \cdots, t$ and $f_4 = \rho - i$. If $m-2-t < \rho \leqq m-2$, the same conclusion holds obviously, with $\rho - i = m-2-t$. The case $\rho = m-1$ is treated later. Hence if $\rho \neq m-1$, there exist integers ζ and σ such that $N = mg + \zeta + \sigma$, $0 \leqq \sigma \leqq m-2-t$, $f_4 = \sigma$, and ζ is one of $\xi, \cdots, \xi + t$. Define $a = 2g + \zeta$, and $b = \zeta$. Then if b and a satisfy (4) the proof is finished. But if b and a satisfy one of (5), then there are exhibited integers g', ζ', σ' such that $N = mg' + \zeta' + \sigma'$, and that $b' = \zeta'$, $a' = 2g' + \zeta'$ do satisfy (4). There are two cases, according as a_5 is not, or is, divisible by t.

First, if $a_5 = 3, 4, 5, 6$, or 8, and if a_5 appears explicitly in σ, then $N = mg + \zeta + a_5 + \sigma - a_5$; and $\zeta' = \zeta + a_5$ and $g' = g$ yield a', b' which do satisfy (4), because $\zeta' \not\equiv \zeta \equiv 0 \pmod t$. Again, if $a_5 = 3, 4, 5, 6$, or 8, and if a_5 does not appear explicitly in σ, then $N = mg + \zeta - a_5 + \sigma + a_5$; and $\zeta' = \zeta - a_5$ and $g' = g$ yield a', b' which do satisfy (4). In each case $r = \sigma'$. There remains the case $a_5 = 7$, for which the proof is long and intricate. Therefore the proof for $\rho = m-1$ is next presented. Here

$$N = mg + \xi + m - 1 = mg + \zeta + (m-2-t),$$

with $\zeta = \xi + t + 1$ and $\sigma = m-2-t$. Then if a, b satisfy (5), $N = m(g+1) + \zeta - t - 2$ yields $g' = g+1$, $\zeta' = \xi - 1$, $\sigma' = 0$; so that $b' = \xi - 1$ and $a' = 2g' + b'$ satisfy (4), since $\zeta' \not\equiv \zeta \equiv 0 \pmod t$.

To complete the proof of Lemma 2 for the case $a_5 = 7$ the following Lemmas 3, 4, 5 and 6 were proved.

LEMMA 3. *Let a, b satisfy* (5_1). *Then* $a + 7$, $b + 7$ *satisfy* (6) *or* (7), *and hence* (4), *if and only if* $h > 1$. *Also* $a - 7$, $b - 7$ *satisfy* (6) *or* (7) *if and only if* $h > 1$, *or* $h = 1$ *and* $A \not\equiv -3 \pmod 7$.

LEMMA 4. *Let a, b satisfy* (5_1), *with* $h = 1$. *Then* $a + 14$, $b + 14$ *satisfy* (6) *or* (7) *if and only if* $A \not\equiv -3 \pmod 7$. *If* $A \equiv -3 \pmod 7$ *then each of the pairs* $a - 14$, $b - 14$, $a + 21$, $b + 21$, $a - 21$, $b - 21$ *satisfies* (6) *or* (7).

LEMMA 5. *Let a, b satisfy* (5_2). *Then* $a + 7$, $b + 7$ *satisfy* (6) *or* (7) *except in the following six cases, when they satisfy* (5); *then the pairs involving* $b + 14$, $b - 14$, $b + 21$, $b + 28$ *satisfy* (6), (7) *or* (5) *as indicated:*

$h = 1$, $B \equiv 1$, $A \equiv B^2 - 3$, $b + 14$ *satisfy* (6) *or* (7),
$h = 1$, $B \equiv -1$, $A \equiv B^2 - 5$, $b - 7$, $b - 14$, $b + 28$ *satisfy* (6) *or* (7); $b + 14$, $b + 21$ *satisfy* (5),
$h = 1$, $B \equiv 2$, $A \equiv B^2 - 6$, $b + 14$ *satisfy* (6) *or* (7),
$h = 1$, $B \equiv -2$, $A \equiv B^2 - 3$, $b + 14$ *satisfy* (6) *or* (7),
$h = 1$, $B \equiv 3$, $A \equiv B^2 - 6$, $b + 21$, $b - 14$ *satisfy* (6) *or* (7); $b + 14$ *satisfy* (5),
$h = 1$, $B \equiv -3$, $A \equiv B^2 - 5$, $b + 14$ *satisfy* (6) *or* (7).

LEMMA 6. *Let a, b satisfy* (5_2). *Then* $a - 7$, $b - 7$ *satisfy* (6) *or* (7) *except in the following eight cases, when they satisfy* (5); *then the pairs involving* $b + 14$, $b - 14$, $b + 21$, $b - 21$, $b + 7$, $b - 28$ *satisfy* (6), (7) *or* (5) *as indicated:*

$h = 1$, $B \equiv 1$, $A \equiv B^2 - 3$, $b + 14$, $b \pm 21$ *satisfy* (6) *or* (7); $b - 14$, (5),
$h = 1$, $B \equiv 1$, $A \equiv B^2 - 5$, $b - 14$, (6) *or* (7),
$h = 1$, $B \equiv 1$, $A \equiv B^2 - 6$, $b - 14$, (6) *or* (7),
$h = 1$, $B \equiv -1$, $A \equiv B^2 - 6$, $b - 14$, (6) *or* (7),
$h = 1$, $B \equiv 2$, $A \equiv B^2 - 5$, $b - 28$, (6) *or* (7); $b - 14$, $b - 21$, (5),
$h = 1$, $B \equiv -2$, $A \equiv B^2 - 6$, $b - 21$, (6) *or* (7); $b - 14$, (5),
$h = 1$, $B \equiv 3$, $A \equiv B^2 - 3$, $b - 14$, (6) *or* (7),
$h = 1$, $B \equiv -3$, $A \equiv B^2 - 5$, $b - 14$, $b + 14$, (6) *or* (7).

The proofs of Lemmas 3, 4, 5, 6 are omitted since they are simple but

intricate. Next note that if $a_5 = 7$ and $a_6 \not\equiv 0 \pmod 7$ or if $a_5 = a_6 = 7$ and $a_7 \not\equiv 0 \pmod 7$, then the proof applied when $a_5 \not\equiv 0 \pmod 7$ is valid, with a_5 replaced by a_J ($= a_6$ or a_7 respectively). There remain three cases with $a_5 = 7$: case I, $(1, 1, 2, 3, 7, 14, \cdots)$ with $n \geqq 7$ and $(1, 1, 2, 3, 7, 7, 21, \cdots)$ with $n \geqq 7$ and $(1, 1, 2, 3, 7, 7, 14, \cdots)$ with $n \geqq 7$; case II, $(1, 1, 2, 3, 7, 7, 7)$; case III, $(1, 1, 2, 3, 7, 7)$. Note that $(1, 1, 2, 3, 7)$ and $(1, 1, 2, 3, 7, 14)$ do not satisfy the necessary conditions (stated later).

Proof for case I. First, if a, b satisfy (5_1) and if a_5 is explicitly in σ: if $h > 1$, use $b' = b + 7$, $\sigma' = \sigma - 7$; but if $h = 1$ and $A \not\equiv -3 \pmod 7$, by Lemmas 3 and 4, use $b' = b + 14$ if also a_6 is explicitly in σ, but use $b' = b - 7$ if a_6 is not explicitly in σ; if $h = 1$ and $A = -3 \pmod 7$, similarly one of $b + 21$, $b \pm 14$, $b - 7$ will be satisfactory. Second, if a, b satisfy (5_1) and if a_5 is not explicitly in σ: by Lemmas 3 and 4, similarly one of $b - 7$, $b + 7$, $b - 14$, $b \pm 21$, will be satisfactory. Next, if a, b satisfy (5_2), by Lemmas 5 and 6 and a lengthy tabulation it was found that it was necessary and sufficient to permit b' to have the values $b \pm 7$, $b \pm 14$, $b \pm 21$, $b \pm 28$; here $g' = g$; also, if $b' = b + \delta$, where δ is one of these multiples of 7, then $a' = 2g' + b'$.

Proof for cases II and III, and for $(1, 1, 2, 3)$. The details for case I were applicable except in certain vital places, where they failed. An independent direct proof, using frequently the device noted in the treatment of $\rho = m - 1$ preceding the statement of Lemma 3, yielded the results tabulated below for b'; always $a' = 2g' + b'$.

$(1,1,2,3,7,7,7)$	b;	$b + 7, b - 9, b - 16, b + 16; b - 7, b + 16, b + 9, b - 16,$
$(1,1,2,3,7,7)$	b;	$b + 7, b - 9, b - 16;\ b - 7, b + 9,$
$(1,1,2,3)$	b and $b - 9$.	

This completes the proof of Lemma 2. Clearly the values of b cluster around ξ. In fact, if $a_5 = 3, 4, 5, 6$, or 8, the maximum b is $\xi + (t - 1) + a_5$ and the minimum b is $\xi + 1 - a_5$. For, the values of b for $\rho = m - 1$ are within this range; also when $\zeta = \xi$ or $\xi + t$ then $b = \zeta \not\equiv 0 \pmod 7$. The consecutive integers $\xi + (t - 1) + a_5, \cdots, \xi + 1 - a_5$ are in number $2a_5 + t - 1$. But in Lemma 2 ξ was an arbitrary integer not divisible by t and $\xi \geqq 4t = 28$. Therefore any set of consecutive integers, $2a_5 + t$ in number, whose smallest integer is greater than 20, can serve as values of b in Lemma 2. Define d as the number of consecutive integers in such a set. Therefore, if $a_5 = 3, 4, 5, 6$, or 8, we have $d = 2a_5 + t$. Similarly for case I of $a_5 = 7$ the maximum and minimum values of b are $\xi + (t - 1) + 28$ and $\xi + 1 - 28$, so that $d = 9t = (t + 2)t$; for case II, they are $\xi + (t - 1) + 16$

and $\xi+1-16$, and $d=6t-3$; for $(1,1,2,3,7,7)$ they are $\xi+(t-1)+9$ and $\xi+1-16$, and $d=4t+4$; for $(1,1,2,3)$ they are $\xi+8$ and $\xi-8$, and $d=18$.

LEMMA 7. *In Lemma 2, any set of d consecutive integers, each $\geqq 21$, can serve as values of b, if $d=2a_5+t$ when $a_5=3,4,5,6$, or 8; $d=(t+2)t$ if $a_5=7$,* case I; $d=6t-3$, case II; $d=4t+4$, case III; $d=18$ *if* $(1,1,2,3)$; $d=2a_J+t$, *when* $a_5=7$, $a_J=a_6\not\equiv 0$ (*mod* 7) *and when* $a_5=a_6=7$; $a_J=a_7\not\equiv 0$ (mod 7).

The final step leading to the desired integers a, b, r mentioned preceding (3) is to choose ξ in Lemma 2 so that (3) is satisfied. This determines the integer M mentioned following (4). The method is essentially that given by Dickson [8] in his improved proof of the Fermat theorem on polygonal numbers, and used in my papers I and II. Since the proof is long and intricate, it is omitted here, except the statement that by the equation $N=r+b+m(a-b)/2$ of Lemma 2 the inequalities (3) are transformed to involve N and m instead of a. The result is that if

$$N\geqq(13d^2-65d+87)m+2(13d-83),$$

and if d is defined as in Lemma 7, and if $N=r+b+m(a-b)/2$, then there are d positive integers within the limits on b which involve N and m instead of a. Since these limits on b imply the limits (3) on b, if ξ is chosen appropriately within these limits then Lemma 2 holds and (3) hold.

The value for d of case I, $a_5=7$ can be lowered if the function is such that there is a coefficient $a_J\not\equiv 0 \pmod 7$ and $a_J<28$. For then the argument applied to $a_5\not\equiv 0 \pmod 7$ in the proof of Lemma 2 holds for a_J. Hence $2a_J+t$ can be used as d. Here $2a_J+t<(t+2)t$ since $a_J<4t=28$.

In section 4 of II necessary and sufficient conditions were found that the functions $(1,1,2,3,\cdots)$ represent the integers $0,1,2,\cdots,34m-16$. They are that $w=m-2$, that $a_k\leqq w_{k-1}+1$ $(5\leqq k\leqq n)$, and that the function be one of the following:

(8)

$(1,1,2,3)$;

$(1,1,2,3,\cdots)$, $a_5=3,4,5,6$, $n\geqq 5$;

$(1,1,2,3,7,\cdots)$, with $a_6=8,\cdots,13$, $n\geqq 6$; also with $a_6=7$, and $n=6$ or $a_7=7,\cdots,14$; also with $a_6=7$, $a_7=15,\cdots,22$ subject to the following conditions (i) or (iii); also with $a_6=14$, $a_7=14,\cdots,28$, $n\geqq 7$;

$(1,1,2,3,8,\cdots)$, with $n\geqq 5$, subject to the following conditions (i) or (ii);

[8] *Bulletin of the American Mathematical Society*, vol. 33 (1927), p. 715.

(i) $a_k \neq w_{k-1}$ for every $k \geqq 6$;

(ii) $a_k = w_{k-1}$ for at least one $k \geqq 6$, and for every such k there is a coefficient a_K satisfying $a_k < a_K \leqq a_k + 7$;

(iii) $a_k = w_{k-1}$ for at least one $k \geqq 6$, and for every such k there is a coefficient a_K satisfying $a_k < a_K \leqq a_k + 14$ but $a_K \neq a_k + 7$.

THEOREM 1. *Let f satisfy the necessary conditions above. Define $M = (13d^2 - 65d + 87)m + 2(13d - 83)$. Then f is universal except perhaps for integers N such that $34m - 16 < N < M$.*

3. *The functions* $(1, 1, 1, 2, \cdots)$. The fundamental structure of the argument is precisely that of paragraph 2. Hence only new or difficult details are given. Sufficient conditions that integers x, y, z, w, each $\geqq 0$, satisfying $a = x^2 + y^2 + z^2 + 2w^2$ and $b = x + y + z + 2w$, exist are (3) and

$$4(5a - b^2) \neq 25^s(5n + e) \qquad (e = 2, 3). \tag{9}$$

These are the conditions of Dickson, since $3 \equiv -2 \pmod 5$. Here $t = 5$.

The conditions which are equivalent to (9) when $a \equiv b \equiv 0 \pmod 5$ are those obtained from (6) and (7) by replacing 7 by 5 and using $e = 2, 3$, and $AB \not\equiv 0 \pmod 5$. A lemma analogous to Lemma 1 holds when $t = 5$, and also one analogous to Lemma 2. If $a_5 = 2, 3, 4, 6$, that is, if $a_5 \not\equiv 0 \pmod 5$, the details are precisely similar to those for $t = 7$ and $a_5 \not\equiv 0 \pmod 7$. The same is true if $a_5 = 5 \not\equiv a_6 \pmod 5$. If $a_5 = 5 \equiv a_6 \pmod 5$, lemmas concerning the pairs involving $b \pm 5$, $b \pm 10$, $b \pm 15$ were proved. These lemmas were distinctly different in statement and proof for the cases $t = 7$ and $t = 5$. The proof was completed as for $t = 7$, and there emerged the following values for d.

LEMMA 8. *Let f satisfy conditions* (12) *of* II (*necessary conditions*). *Let $d = 2a_5 + t$ if $a_5 = 2, 3, 4, 6$; $d = 2a_6 + t$ if $a_5 = 5 \not\equiv a_6 \pmod 5$; $d = (t + 2)t$ if $a_5 = 5 \equiv a_6 \pmod 5$ and $n > 5$ except $d = 27$ if $(1, 1, 1, 2, 5, 5)$; $d = 14$ if $n = 4$. Let $N \geqq \xi \geqq 3t$. Then there are integers a, b, r, each $\geqq 0$, such that $N = r + b + m(a - b)/2$, $a \equiv b \pmod 2$,* (9) *is satisfied, and $f_4 = r \leqq m - 2 - t$. Any set of d consecutive integers, each $\geqq 3t$, can serve as values of b.*

The value $(t + 2)t$ for d can be lowered if there is a coefficient $a_J \not\equiv 0 \pmod t$ and $a_J < 3t$. For the argument applied to $a_5 \not\equiv 0 \pmod t$ holds for a_J. Hence $2a_J + t$ can be used as d, and $2a_J + t < (t + 2)t$.

In section 4 of II necessary and sufficient conditions were found that the functions $(1, 1, 1, 2, \cdots)$ represent the integers $0, 1, \cdots, 34m - 16$. They are (12) of II.

THEOREM 2. *Let f satisfy* (12) *of* II. *Define*

$$M = (9d^2 - 45d + 61)m + 2(9d - 49).$$

Then f represents every positive integer N except perhaps $34m - 16 < N < M$.

The results for $(1, 1, 1, 2, \cdots)$ of Theorem 3 of II are improved upon in this theorem.

4. *The functions* $(1, 1, 2, 4, \cdots)$. The fundamental structure of the argument is that of paragraph 2. Dickson proved that if b is odd or double an odd integer, and if (3) hold, then

$$a = x^2 + y^2 + 2z^2 + 4w^2, \qquad b = x + y + 2z + 4w \tag{10}$$

have a solution in integers $\geqq 0$. The new case in which b is divisible by 4 is included in the following theorem proved in section 7:

THEOREM 3. *Let $b = 2^iB$, $a = 2^{2h-1}A$ or $2^{2h}A$, where i and h are integers $\geqq 1$, while A and B are odd integers. Then* (10) *have solutions in integers $\geqq 0$ if and only if $8a \geqq b^2$ and one of the following conditions hold:*

$$\begin{cases} a = 2^{2h}A, & 1 \leqq i \leqq h + 1, \quad 1 \leqq h; \\ a = 2^{2h}A, & 3 \leqq h + 2 = i, \quad A \not\equiv 1 \pmod 8; \\ a = 2^{2h}A, & 4 \leqq h + 3 \leqq i, \quad A \not\equiv 7 \pmod 8; \end{cases} \tag{11}$$

$$\begin{cases} a = 2^{2h-1}A, & 2 < i = h + 1, \quad B \equiv \pm 1 \pmod 8, \quad A \not\equiv 15 \pmod{16}; \\ a = 2^{2h-1}A, & 2 < i = h + 1, \quad B \equiv \pm 3 \pmod 8, \quad A \not\equiv 7 \pmod{16}; \\ a = 2^{2h-1}A, & 2 = i = h + 1, \quad A \equiv \pm 3 \pmod 8; \\ a = 2^{2h-1}A, & 1 \leqq i \neq h + 1, \quad 1 \leqq h. \end{cases} \tag{12}$$

Lemma 1 holds with $t = 8$. In paragraph 4 of II necessary and sufficient conditions were found that the functions $(1, 1, 2, 4, \cdots)$ represent the integers $0, 1, \cdots, 34m - 16$, namely merely that $w = m - 2$ and that $n = 4$ or that $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$. A lemma analogous to Lemma 2 holds: if $a_5 \not\equiv 0 \pmod 4$, that is, if $a_5 = 5, 6, 7, 9$ the details are similar to those for $t = 5, 7$ and $a_5 \not\equiv 0 \pmod t$; but if $a_5 = 4, 8$ the supplementary lemmas

concerning the pairs involving $b \pm 4$, $b \pm 8$, $b \pm 12$, $b \pm 16$ were extremely intricate. The proof was then completed as for $t = 5, 7$ and the values of d determined.

LEMMA 9. *Let f satisfy $w = m - 2$, and $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$ or $n = 4$. Let $d = 2a_5 + t$ if $a_5 = 5, 7, 9$; $d = 2a_6 + t$ if $a_5 \equiv 0 \not\equiv a_6 \pmod 4$; $d = 5t$ if $a_5 = 8$ and another coefficient is 8 or 16; $d = 4t$ if $a_5 = 8$ and another coefficient is 12; $d = 28$ if $(1, 1, 2, 4, 8)$; $d = 5t/2$ if $a_5 = 4$ and another coefficient is 4 or 8; $d = (t-1)t/2$ if $a_5 = 4$ and another coefficient is 12; $d = 3t$ if f is $(1, 1, 2, 4, 4)$; $d = 20$ if $(1, 1, 2, 4)$.* Let $N \geqq \xi \geqq 2t$. *Then there are integers a, b, r, each $\geqq 0$, such that $N = r + b + m(a-b)/2$, $a \equiv b \equiv 1 \pmod 2$ or (11) or (12) is satisfied, and $f_4 = r \leqq m - 2 - t$. Any set of d consecutive integers, each $\geqq 2t$, can serve as values of b.*

The value for d can be lowered if $a_5 \equiv 0 \pmod 4$ but there exists a coefficient $a_J \not\equiv 0 \pmod 4$ such that $2a_J + t$ is less than the value stated in Lemma 9.

THEOREM 4. *Let f satisfy $w = m - 2$, and $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$ or have $n = 4$. Define $M = (15d^2 - 75d + 100)m + 2(15d - 103)$. Then f represents every positive integer except perhaps $34m - 16 < N < M$.*

5. *The functions* $(1, 1, 1, 1, \cdots)$. Again the fundamental structure of the argument is that of paragraph 2. Sufficient conditions that integers x, y, z, w, each $\geqq 0$, satisfying $a = x^2 + y^2 + z^2 + w^2$, $b = x + y + z + w$, exist are (3) and

$$4a - b^2 \neq 4^s(8n + 7). \tag{13}$$

When $a \equiv 0 \equiv b \pmod 2$ the conditions, equivalent to (13), suitable to deriving supplementary lemmas for pairs involving $b \pm 4$, $b \pm 8$, are (14) and (15), using $a = 2^{2h}A$ or $2^{2h-1}A$, $b = 2^iB$, with $AB \not\equiv 0 \pmod 2$:

$$\begin{aligned} &a = 2^{2h}A, \quad 1 \leqq i = h \text{ or } h + 1; \text{ or} \\ &a = 2^{2h}A, \quad 3 \leqq h + 2 = i, \quad A \not\equiv 3 \pmod 8; \text{ or} \\ &a = 2^{2h}A, \quad 4 \leqq h + 3 \leqq i, \quad A \not\equiv 7 \pmod 8; \end{aligned} \tag{14}$$

$$a = 2^{2h-1}A, \quad 1 \leqq h \leqq i. \tag{15}$$

Lemma 1 holds with $t = 4$. In paragraph 4 of II the necessary and sufficient conditions that the functions $(1, 1, 1, 1, \cdots)$ represent the integers

$0, 1, \cdots, 34m - 16$ were merely that $w = m - 2$ and that $n = 4$ or that $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$. A lemma analogous to Lemma 2 holds. The previous treatment of $\rho = m - 1$ is invalid here, because, if a is odd or double an odd, then a, b satisfy (15), but if $a \equiv 0 \pmod 4$, then when

$$b = \zeta = \xi + t + 1, \qquad a = 2g + \zeta$$

fail to satisfy (14) or (15) so also do $b' = \zeta - t - 2$, $a' = 2(g + 1) + b'$. If $a_5 = 1, 2, 3, 5$ then one of ζ or $\zeta' = \zeta \pm a_5$ is satisfactory; similarly if $a_5 = 4 \not\equiv a_6 \pmod 4$; but if $a_5 = 4 \equiv a_6 \pmod 4$ then one of ζ, $\zeta \pm 4$, or $\zeta \pm 8$ is satisfactory (in fact ζ or $\zeta \pm a_J$, if there exists $a_J \not\equiv 0 \pmod 4$ such that $a_J < 8$), if $n > 5$, while one of ζ, $\zeta \pm 4$, $\zeta + 10$ is satisfactory if $n = 5$. The details for $\rho < m - 1$ are similar to those for $t = 5, 7, 8$, if $a_5 = 1, 2, 3, 5$; if $a_5 = 4$ supplementary lemmas involving $b \pm 4$, $b \pm 8$ were used. The proof was completed then, and the values of d determined.

If $n = 4$ the function is known to be universal.[9]

LEMMA 10. *Let f satisfy $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$. Let $d = 2a_5 + 6$ if $a_5 = 1, 2, 3, 5$; $d = 2a_6 + 6$ if $a_5 = 4 \not\equiv a_6 \pmod 4$; $d = 22$ if $a_5 = 4$ and another coefficient is t or $2t$; $d = 5t$ for $(1, 1, 1, 1, 4)$. Let $N \geqq \xi \geqq t$. Then there are integers a, b, r, each $\geqq 0$, such that $N = r + b + m(a - b)/2$, $a \equiv b \pmod 2$, (13) holds, and $f_4 = r \leqq m - 2 - t$. Any set of d consecutive integers, each $\geqq t$, can serve as values of b.*

The value $d = 5t + 2$, when $a_5 = t$, can be lowered if there exists a coefficient $a_J \not\equiv 0 \pmod 4$ such that $a_J < 2t$, for then the argument applied to $a_5 \not\equiv 0 \pmod 4$ is valid, and $d = 2a_J + 6$.

THEOREM 5. *Let f have $w = m - 2$, and $a_k \leqq w_{k-1} + 1$ $(5 \leqq k \leqq n)$. Define $M = (7d^2 - 35d + 48)m + 2(7d - 35)$. Then f represents every positive integer N except perhaps $34m - 16 < N < M$.*

6. *The functions* $(1, 1, 2, 2, \cdots)$. The fundamental structure of the argument is that of paragraph 2. The method is also that which yielded Lemma 7 of II. The results of Lemma 7 of II (in which the upper value $\beta + 5 + a_5$, if $a_5 = 3, 5, 7$ should be $\beta + 7 + a_5$) are not as good as the new results here obtained, because my limits on b in Lemma 6 of II are not as good as (3) which Dickson obtained, and because of improved values for d. If $a_5 = 3, 4, 5, 7$ then for $\rho = m - 1$ satisfactory values of b are ζ or $\zeta + a_5$;

[9] Dickson, *Journal de Mathématiques*, ser. 9, vol. 7 (1928), Theorems 11-15.

if $a_5 = 2, 6$ then ζ, $\zeta + a_5$ or $\zeta - t - 2$ are satisfactory. For $\rho < m - 1$ and $a_5 = 3, 4, 5, 7$ the details are similar to those for $t = 7$ and $a_5 \not\equiv 0 \pmod 7$; for $a_5 = 2, 6$ supplementary lemmas involving $b \pm 2$, $b \pm 4$, $b \pm 6$, $b \pm 8$, $b \pm 10$, $b \pm 12$ were used. The details were different for $a_5 = 2$ and $a_5 = 6$, but a striking uniformity of values for d emerged.

The necessary conditions (13) of II should have been stated so that $(1, 1, 2, 2, 6)$ is excluded.

LEMMA 11. *Let f satisfy* (13) *of* II *excluding* $(1, 1, 2, 2, 6)$. *Let $d = 2a_5 + t + 2$ if $a_5 = 3, 4, 5, 7$; if $a_5 = 2$ or 6 and $a_6 > a_5 + 1$, let $d = 2a_6 + t$; if $a_5 = 2$ or 6 and $a_6 = a_5 + 1$, let $d = 2a_6 + t + 1$;* if $a_5 = 2$ *or 6 and $a_6 = a_5$, let $d = 4a_5 + t$; for $(1, 1, 2, 2, 2)$ and $(1, 1, 2, 2, 6)$ let $d = 2(t + 2) + a_5$; for $(1, 1, 2, 2)$ let $d = 25$. Let $N \geqq \xi \geqq 2t$. Then there are integers a, b, r, each $\geqq 0$, such that $N = r + b + m(a - b)/2$, the equations $a = x^2 + y^2 + 2z^2 + 2w^2$, $b = x + y + 2z + 2w$ have solutions in integers each $\geqq 0$, and $f_4 = r \leqq m - 2 - t$. Any set of d consecutive integers, each $\geqq 2t$, can serve as values of b.*

The values of d can be lowered, if $a_5 = 2$ or 6 and $a_6 = a_5$, to $2a_J + t + 1$ if $a_J = a_5 + 1$ and to $2a_J + t$ if $a_5 + 1 < a_J < 2a_5$.

THEOREM 6. *Let f satisfy* (13) *of* II *excluding* $(1, 1, 2, 2, 6)$. *Define $M = (11d^2 - 55d + 74)m + 2(11d - 65)$. Then f represents every positive integer N except perhaps $34m - 16 < N < M$.*

7. *Solvability of* (10). Dickson showed that necessary conditions are

$$a \equiv b \pmod 2, \tag{16}$$

$$8a - b^2 = F^2 + 2v^2 + 4W^2 \tag{17}$$

$$F = 4w - 2z - y - x, \qquad v = 2z - y - x, \qquad W = y - x. \tag{18}$$

The solution of (18) with (10_2) give

$$8w = b + F, \qquad 8z = b - F + 2v, \qquad 8y = b - F - 2v + 4W, \qquad 8x = b - F - 2v - 4W. \tag{19}$$

Let a and b be even, expressed as in Theorem 3. By examining the conditions for equality, we find the inequality $8a - b^2 \neq 4^m(16n + 14)$ holds only in the following cases:

$$(20)\quad \begin{cases} a = 2^{2h-1}A, & h+1=i, \quad B \equiv \pm 1 \pmod 8, \quad A \not\equiv 15 \pmod{16}; \\ a = 2^{2h-1}A, & h+1=i, \quad B \equiv \pm 3 \pmod 8, \quad A \not\equiv 7 \pmod{16}; \\ a = 2^{2h-1}A, & h+1 \neq i; \end{cases}$$

$$(21)\quad \begin{cases} a = 2^{2h}A, & h+1 \geqq i; \\ a = 2^{2h}A, & h+2=i, \quad A \not\equiv 1 \pmod 8; \\ a = 2^{2h}A, & h+3 \leqq i, \quad A \not\equiv 7 \pmod 8. \end{cases}$$

Use is made of the known

LEMMA 12. *If an even integer is represented by $x^2+y^2+2z^2$, there exists a representation with x and y both even.*

Let a and b satisfy (20) or (21) with $i \geqq 2$. Then $8a - b^2 \equiv 0 \pmod 4$, and in $8a - b^2 = F^2 + 2v^2 + 4W^2$ we have F and v both even. Hence there are integers F_1 and v_1 such that $F = 2F_1$ and $v = 2v_1$ and (22) or (23) holds, according as $a = 2^{2h-1}A$ or 2^{2h}:

$$(22)\qquad 2^{2h}A - 2^{2i-2}B^2 = F_1{}^2 + 2v_1{}^2 + W^2,$$

$$(23)\qquad 2^{2h+1}A - 2^{2i-2}B^2 = F_1{}^2 + 2v_1{}^2 + W^2.$$

Since $i \geqq 2$, the left member of (22) and that of (23) are divisible by 4. Hence by Lemma 12 we can take F_1 and W both even; that is, there exist integers F_2, W_2 and v_2 such that $F_1 = 2F_2$, $W = 2W_2$ and $v_1 = 2v_2$, and that (22) and (23) become respectively

$$(24)\qquad 2^{2h-2}A - 2^{2i-4}B^2 = F_2{}^2 + 2v_2{}^2 + W_2{}^2,$$

$$(25)\qquad 2^{2h-1}A - 2^{2i-4}B^2 = F_2{}^2 + 2v_2{}^2 + W_2{}^2.$$

By these values $F = 4F_2$, $v = 4v_2$ and $W = 2W_2$, (19) become

$$(26)\quad \begin{aligned} 2w &= 2^{i-2}B + F_2, & 2z &= 2^{i-2}B - F_2 + 2v_2, \\ 2y &= 2^{i-2}B - F_2 - 2v_2 + 2W_2, & 2x &= 2^{i-2}B - F_2 - 2v_2 - 2W_2. \end{aligned}$$

If $i \geqq 3$ conditions (26) are equivalent to F_2 even, but if $i = 2$ they are equivalent to F_2 odd, since B is odd. Now, if $i \geqq 3$ by Lemma 12 we can take F_2 even in (25); also in (24), if $i \geqq 3$ and $h > 1$. If $i \geqq 3$ and $h = 1$ in (24), then one of F_2 and W_2 is odd and the other is even, since A is odd, and hence by the symmetry we can take F_2 even. If $i = 2$ in (25), then $B + F_2 + W_2$ is even and hence $F_2 \not\equiv W_2 \pmod 2$, and we can take F_2 odd by the symmetry. If $i = 2$ and $h > 1$ in (24) the same argument holds. There remains therefore the case $i = 2$, $h = 1$ in (24); but in (24), as in

(22), h and i satisfy (20), and hence in fact (20_1) or (20_2). Here (24) becomes $A - B^2 = F_2{}^2 + 2v_2{}^2 + W_2{}^2$.

Lemma 13 states necessary and sufficient conditions on a and b, satisfying (20_1) or (20_2) with $h = 1$, that there exist a representation of

$$A - B^2 = F_2{}^2 + 2v_2{}^2 + W_2{}^2$$

with F_2 odd.

LEMMA 13. *If A and B satisfy $B \equiv \pm 1 \ (mod\ 8)$ and $A \not\equiv 15 \ (mod\ 16)$, or $B \equiv \pm 3 \ (mod\ 8)$ and $A \not\equiv 7 \ (mod\ 16)$, if A is odd, and if $A - B^2$ is represented by the form $x^2 + y^2 + 2z^2$, then there exists a representation with x odd if and only if $A \equiv \pm 3 \ (mod\ 8)$.*

First, let $B \equiv \pm 1 \pmod 8$. If $A \equiv 7 \pmod{16}$ then $A - B^2 \equiv 6 \pmod{16}$. By the proof of Lemma 12, we have $(A - B^2)/2 = s^2 + t^2 + z^2$, and therefore $s \equiv t \equiv z \equiv 1 \pmod 2$, $x \equiv y \equiv 0 \pmod 2$. Similarly, if $A \equiv 1$ or $9 \pmod{16}$, we have $s \equiv t \equiv z \equiv 0 \pmod 2$, $x \equiv y \equiv 0 \pmod 2$. Next let $A \equiv 3, 5, 11$, or $13 \pmod{16}$. Then if in $A - B^2 = x^2 + y^2 + 2z^2$ in fact x and y are even, we have $s \equiv t \pmod 2$. If $z \not\equiv s \pmod 2$, we have

$$(A - B^2)/2 = s^2 + t^2 + z^2, \qquad A - B^2 = (s + z)^2 + (s - z)^2 + 2t^2$$

with $s + z$ odd. But if $z \equiv s \pmod 2$ then in fact each of z, s, t is even or each is odd. If each is even, then $(A - B^2)/2$ is divisible by 4 and therefore $A - B^2$ is divisible by 8; but by the hypotheses $B \equiv \pm 1 \pmod 8$ and $A \equiv 3, 5, 11, 13 \pmod{16}$, $A - B^2 \not\equiv 0 \pmod 8$. Similarly a contradiction is obtained if $z \equiv s \equiv t \equiv 1 \pmod 2$. This completes the proof when $B \equiv \pm 1 \pmod 8$.

Next, if $B \equiv \pm 3 \pmod 8$, it is shown similarly that $A \equiv 1, 9, 15 \pmod{16}$ imply $x \equiv y \equiv 0 \pmod 2$. But if $A \equiv 3, 5, 11, 13 \pmod{16}$, then $x \equiv y \equiv 0 \pmod 2$ imply $z \equiv s \pmod 2$ and

$$A - B^2 = (s + z)^2 + (s - z)^2 + 2t^2$$

with $s + z$ odd. This completes the proof of Lemma 13.

Therefore if a and b satisfy (20) with $i = 2$ and $h = 1$, there exists a representation (17) such that (19) yield integers if and only if $A \equiv \pm 3 \pmod 8$.

This proves Theorem 3.

Since (11) and (12) are merely (20) and (21) with the case $i = 2$, $h = 1$ of (20) modified to include only $A \equiv \pm 3 \pmod 8$, an alternative statement to Theorem 3 is

THEOREM 7. *If a and b are even integers such that*

$$8a - b^2 \neq 4^m(16n + 14)$$

and that $8a \geqq b^2$ then there exist integers x, y, z, w satisfying (10) *if and only if $b \not\equiv 4 \pmod 8$, or $a \equiv 0 \pmod 4$, or $a \equiv \pm 6 \pmod{16}$.*

Dickson proved that if k is an integer $\geqq 0$ and if (17) holds, then the values x, y, z, w from (19) are each $> -k$ if $7a < b^2 + 2bk + 8k^2$. The proof does not depend upon whether b is divisible by four or not.

THEOREM 8. *Let k be an integer $\geqq 0$, and let a and b be integers such that $a \equiv b \pmod 2$, $8a \geqq b^2$, $b \geqq 8(1 - k)$ and $7a < b^2 + 2bk + 8k^2$. Then there exist integers x, y, z, w, each $> -k$, satisfying* (10) *if and only if $b \not\equiv 4 \pmod 8$, or $a \equiv 0 \pmod 4$, or $a \equiv \pm 6 \pmod{16}$, and $8a - b^2 \neq 4m(16n + 14)$.*

NORTHWESTERN UNIVERSITY,
EVANSTON, ILL.

A CUBIC ANALOGUE OF THE CAUCHY-FERMAT THEOREM.[1]

By Alvin Sugar.

Introduction. In this paper we shall obtain an ideal universal Waring theorem for the polynomial

$$(1) \qquad P(x) = m(x^3 - x)/6 + x, \qquad x \text{ integral and } \geqq 0,$$

where m is an integer $\geqq 16$, i. e. we shall prove $g(P) = m + 3$ for $m \geqq 16$.

In Part II of this paper we evaluate a constant $C_1 = 10^{12}m^{10}$, which maximizes the constants of papers of Dickson,[2] Baker[3] and Webber[4]; this gives us the following theorem.

THEOREM 1. *For $m \geqq 7$, every integer $\geqq C_1 = 10^{12}m^{10}$ is a sum of nine or ten values of* (1) *according as the congruence* $m \equiv 6 \pmod 9$ *does not or does hold.*

In Part I we develop a powerful ascension theorem and ascension methods, and by ascending beyond the constant C_1, prove that every positive integer $\leqq C_1$ is a sum of $m + 3$ values of $P(x)$ for $m \geqq 16$.

PART I. ASCENSION METHODS.

1. *Ascension theorems.* We shall ascend beyond the constant $C_1 = 10^{12}m^{10}$ first for a fixed range and then for an arbitrary range of values of m.

We write

$$F(a) = P(a+1) - P(a)$$

and apply a theorem of Dickson's[5] to our polynomial $P(x)$.

THEOREM 2. *Let every integer n, $c < n \leqq g$, be a sum of $k - 1$ values of $P(x)$, and let a be an integer $\geqq 0$ for which $F(a) < g - c$. Then every integer N, $c < N \leqq g + P(a+1)$, is a sum of k values of $P(x)$.*

---

[1] Presented to the Society, November 30, 1935.

[2] *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 1-12.

[3] Doctoral dissertation, Chicago, 1934.

[4] *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 493-510.

[5] Theorem 9 in the *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 709.

Before going on to our next theorem, we note that in this series of theorems, g and c need not be integers.

THEOREM 3. *Let every integer n, $c < n \leqq g$, be a sum of $k-1$ values of $P(x)$, and let y be a real number $\geqq 0$ which satisfies the inequality $F(y+1) < g - c$. Then every integer N,*

$$c \leqq N \leqq g + P(y+1), \tag{2}$$

is a sum of k values of $P(x)$.

By way of proof we observe that $F([y]+1) \leqq F(y+1)$ and $P(y+1) < P([y]+2)$, since $F(x)$ and $P(x)$ are properly monotone increasing functions for $x \geqq 1$.

We are now in a position to introduce an important ascension theorem, which will enable us to breach a huge interval in one step. The inequality

$$F(3t^{(3/2)e}+1) < P(3t^{e}+1) \tag{3}$$

holds for $t \geqq 1$. Let t be a real number $\geqq 1$ which satisfies the inequality $F(3t+1) < pm + q$, and let every integer N_0, $c < N_0 \leqq c + pm + q$, be a sum of k values of $P(x)$. Then from (2) we have that every integer N', $c < N' \leqq c + pm + q + P(3t+1)$, and hence every integer N_1, $c < N_1 \leqq c + P(3t+1)$, is a sum of $k+1$ values. Similarly, since $F(3t^{3/2}+1) < P(3t+1)$, by (3) with $e = 1$, then every integer N_2, $c < N_2 \leqq c + P(3t^{3/2}+1)$, is a sum of $k+2$. And finally, every integer N_2,

$$c < N_2 \leqq c + P(3t^{(3/2)^{s-1}}+1),$$

is a sum of $k+s$. The proof of this last statement is made by an induction on s. Since

$$(9/2)t^{2(3/2)^s}m < c + P(3t^{(3/2)^{s-1}}+1),$$

we may state the following theorem.

THEOREM 4. *Let every integer n, $c < n \leqq c + pm + q$, be a sum of k values of $P(x)$, and let t be a real number $\geqq 1$ which satisfies the inequality $F(3t+1) < pm + q$. Then every integer N,*

$$c < N \leqq (9/2)t^{2(3/2)^s}m$$

is a sum of $k+s$ values of $P(x)$.

2. *The first ascension*; $16 \leqq m \leqq 1950$. There follows a list of values of $P(x)$.

$$0, 1, a = m + 2;\ b = 4m + 3,\ c = 10m + 4,\ d = 20m + 5,\ e = 35m + 6,$$
$$f = 56m + 7,\ g = 84m + 8,\ h = 120m + 9,\ i = 165m + 10.$$

We are also going to list a set of intervals such that an integer lying in anyone of these intervals will be a sum of $m - 8$ values of $P(x)$. These intervals will overlap for $m \geqq 16$. Therefore $m - 8$ values will suffice over the interval defined by the overlapping intervals.

We shall reconstruct a portion of the following list. We begin with $120m + 9$. By adding $m - 8$ to this we obtain $121m + 1$. It is evident that every integer from $120m + 9$ to and not including $121m + 1$ is a sum of $m - 8$ values. Now consider the integer $120m + 16 = a + e + g$. By adding $m - 10$ to this we obtain $121m + 6$. It is evident that every integer from $120m + 16$ to $121m + 6$ is a sum of $m - 8$ values, and continuing thus we come to the interval $(120m + 24, 121m + 11)$ over which $m - 8$ values will suffice. Since $121m + 11 = h + a$, we can begin all over again as we did with $120m + 9$ by adding $m - 9$ to $121m + 11$ and repeating the above procedure. By inspection it may be verified that the following set of intervals overlap for $m \geqq 16$.

$(h = 120m + 9, 121m + 1)$, $(a + e + g = 120m + 16, 121m + 6)$, $(2b + 2f = 120m + 20, 121m + 9)$, $(2a + b + c + d + g = 120m + 24, 121m + 11)$, $(a + h = 121m + 11, 122m + 2)$, $(2a + e + g = 121m + 18, 122m + 7)$, $(c + d + e + f = 121m + 22, 122m + 11)$, $(3c + e + f = 121m + 25, 122m + 13)$, $(2a + h = 122m + 13, 123m + 3)$, $(c + 2f = 122m + 18, 123m + 8)$, $(a + c + d + e + f = 122m + 24, 123m + 12)$, $(a + 3c + e + f = 122m + 27, 123m + 14)$, $(2a + c + 2d + 2e = 122m + 30, 123m + 16)$, $(1 + 3a + h = 123m + 16, 124m + 4)$, $(a + c + 2f = 123m + 20, 124m + 9)$, $(a + 2b + c + d + g = 123m + 25, 124m + 12)$, $(b + h = 124m + 12, 125m + 3)$, $(a + b + e + g = 124m + 19, 125m + 8)$, $(4c + g = 124m + 24, 125m + 12)$, $(2a + 2b + c + d + g = 124m + 27, 125m + 13)$, $(b + c + 2d + 2e = 124m + 29, 125m + 16)$, $(2 + a + b + h = 125m + 16, 126m + 4)$, $(a + 2d + g = 125m + 20, 126m + 9)$, $(b + c + d + e + f = 125m + 25, 126m + 13)$.

Hence if $m \geqq 16$, then every integer n, $120m + 8 < n \leqq 126m + 12$, is a sum of $m - 8$ values of $P(x)$. Applying Theorem 2 we see that $F(3) = 6m + 1 < 6m + 4$; then $m - 7$ values suffice from $120m + 8$ to $120m + 8 + 16m + 8$. Two more applications of this theorem give the result that $m - 5$ values will suffice from $120m + 8$ to $120m + 8 + 216m$.

The next ascent will be made in one step by employing Theorem 4. Since $t = 19/3$ satisfies the inequality $F(3t+1) < 216m$, then every integer N,

$$120m + 8 < N \leqq c_1 m = (9/2)(19/3)^{2(3/2)^8} m,$$

is a sum of $m+3$ values. It is evident that the inequality $10^{12}m^{10} \leqq c_1 m$ holds for $16 \leqq m \leqq 1950$. Hence we have by employing Theorem 1 the following theorem.

THEOREM 5. *Let m have the range $16 \leqq m \leqq 1950$; then every positive integer $> 120m + 8$ is a sum of $m+3$ values of $P(x)$.*

3. *The second ascension.* Again we construct a set of overlapping intervals. This time we begin with an arbitrary value $P(A) = Rm + A$, and we take $m - r$ as the number of values which will suffice over each interval. By adding $m - r$ (where r is a positive integer) to $P(A)$, we obtain the first interval

$$\big(Rm + A, (R+1)m + A - r\big),$$

We take $r = R - A - 10$. The rest of the intervals can be written at once, as follows:

$$\begin{gathered}\big((R-1+t)m + 2R - 2 + 2t, (R+t)m + R + t - r\big),\\ \big((R+t)m + R + t - r, (R+1+t)m + R - 2r + t - 10)\big).\\ (t = 1, \cdots, 10).\end{gathered}$$

We observe that $(R-1+t)m + 2R - 2 + 2t = (R-1+t)a$ and that for this range of t the integer[6] $(R+t)m + R + t - r = P(A) + ta + 10 - t$ is a sum of 11 values of $P(x)$. By inspection it is evident that these intervals will overlap for $m \geqq Q(A) = 3R - 2A - 1 = (A^3 - 5A)/2 - 1$. We also see that $r = (A^3 - 7A)/6 - 10$.

LEMMA 1. *For $m \geqq Q(A)$, $A \geqq 5$, every integer n,*

$$Rm + A \leqq n \leqq (R+10)m + A + 20,$$

is a sum of $m - r$ values of $P(x)$.

In the following discussion we shall prove statements (S_1) and (S_2). We begin with $P(A)$ and show (S_1) *that for $Q(A) \leqq m \leqq Q(A+1)$ every*

[6] As a matter of fact the value assigned to r was obtained by requiring that r be the greatest integer for which the inequality $(R+t)m + R + t - r \geqq P(A) + ta$, $(t = 1, \cdots, 10)$ holds.

integer $> P(A)$ *is a sum of* $m+3$ *values, provided* $A \geqq 10$. We also show (S_2) *that for* $m \geqq Q(A)$, $m+3$ *values will suffice from* $P(A)$ *to* $P(A+1)$ *inclusive, when* $A \geqq 10$. Since $Q(A)$ is an increasing function, then, by (S_2) and an induction on A, we conclude (S_3) *that for* $m \geqq Q(A)$, *every integer* n, $P(10) \leqq n \leqq P(A+1)$, *will be a sum of* $m+3$ *values*. Hence from (S_1) and (S_3) we have the following theorem.

THEOREM 6. *Every integer* $\geqq P(10)$ *is a sum of* $m+3$ *values of* $P(x)$ *for* $m \geqq Q(10) = 474$.

This and Theorem 5 give us the next theorem.

THEOREM 7. *Every integer* $\geqq P(10)$ *is a sum of* $m+3$ *values of* $P(x)$ *for* $m \geqq 16$.

There remains yet to be proved, the Statements (S_1) and (S_2). In establishing these statements we make use of a pair of inequalities which are derived from the expansion of k^x, into the power series

$$k^x = 1 + x \log k + \frac{x^2 \log^2 k}{2!} + \frac{x^3 \log^3 k}{3!} + \cdots.$$

Since this series converges for all x, we have the following inequalities holding for a positive x.

$$k^x > \frac{x^2 \log^2 k}{2}, \qquad k^x > \frac{x^3 \log^3 k}{6}. \tag{4}$$

To the results of Lemma 1 we apply Theorems 2 and 4, and we find that every integer N,

$$P(A) \leqq N \leqq c_2 m = (9/2)(10)^{2(3/2)^r} m, \tag{5}$$

is a sum of $m+3$ values of $P(x)$ for $m \geqq Q(A)$. We know that

$$10^{12} m^{10} \leqq c_2 m, \quad \text{when} \quad Q(A) \leqq m \leqq M = (10^{-12} c_2)^{1/9}.$$

Let $y = A^3/10$; it is then evident that [7]

$$M > (10)^{(2/9)(3/2)^y - 12/9};$$

for,

$$r = (A^3 - 7A)/6 - 10 > A^3/10.$$

From (4_1), we have

$$\frac{2}{9}\left(\frac{3}{2}\right)^y - \frac{12}{9} > \frac{y^2 \log^2 1.5}{9} - \frac{12}{9} > 10^{-4} A^6,$$

[7] For the remainder of this discussion we shall take $A \geqq 10$.

Write $z = 10^{-4}A^6$; then $M > 10^z$. Employing (4_2), we obtain

$$10^z > 10^{-12}A^{18} > A^3 > (A^3 + 3A^2 - 2A - 6)/2 = Q(A+1).$$

Therefore $M > Q(A+1)$, and we have proved (S_1).

It is evident that

$$c_2 m > Mm > m\,Q(A) > P(A+1).$$

This result and (5) prove (S_2).

4. *The positive integers* $< P(10)$. We shall prove another lemma.

LEMMA 2. *Every positive integer* $\leqq c = 10m + 4$ *is a sum of* $m + 3$ *values of* $P(x)$ *for* $m \geqq 4$.

It is evident that every integer $< 3m + 6 = 3a$ is a sum of $m + 3$ values. Adding $m + 1$ to $3m + 6$, we see that every integer $< 4m + 7 = 4 + b$ is a sum of $m + 3$ values. Adding $m - 1$ to this, we see that every integer $< 5m + 6 = 1 + a + b$ is a sum of $m + 3$ values. Repetition of this argument gives the following list:

$$6m + 7 = 2a + b, \quad 7m + 8 = 6a + m - 4, \quad 7m + 10 = 1 + 3a + b,$$
$$8m + 9 = 3 + 2b, \quad 9m + 8 = a + 2b, \quad 10m + 4 = c.$$

This completes the proof of the lemma.

The following set of intervals, which overlap for $m \geqq 16$, give rise to the conclusion that every integer n,

$$(6) \qquad 10m + 3 < n \leqq 13m + 9,$$

is a sum of $m - 1$ values of $P(x)$.

$(c = 10m + 4,\ 11m + 3)$, $(6a + b = 10m + 15,\ 11m + 8)$, $(2 + a + c = 11m + 8,\ 12m + 4)$, $(7a + b = 11m + 17,\ 12m + 9)$, $(1 + 2a + c = 12m + 9,\ 13m + 5)$, $(8a + b = 12m + 19,\ 13m + 10)$.

Applying Theorem 2 to (6) four times, we obtain the following result.

LEMMA 3. *For* $m \geqq 16$, *every integer* n, $10m + 3 < n \leqq 217m + 32$, *is a sum of* $m + 3$ *values of* $P(x)$.

This result along with Lemma 2 and Theorem 7 completes the proof of the Principal Theorem.

PRINCIPAL THEOREM. *Every positive integer is a sum of $m+3$ values of $m(x^3-x)/6+x$ for non-negative integers x, where $m \geqq 16$.*

PART II. EVALUATION OF THE CONSTANT.

The proof of Theorem 2 of Dickson's and of similar theorems of Baker's and Webber's[8] depends upon the existence of an integer C lying in each interval of a triple of intervals of the form

$$(7) \qquad f(m, b_i) \leqq 3^{i-1}C \leqq F(m, b_i) \qquad (i=1,2,3)$$

where the b_i are suitably chosen positive odd integers. For $m \not\equiv 6 \pmod 9$, (7) takes on the form (8). For $m \equiv 6 \pmod 9$, (7) becomes (9).

$$(8) \qquad 2m+\gamma+(9/8)mb_i^3 \leqq 3^{i-1}C \leqq (3/2)mb_i^3+m/3 \qquad (i=1,2,3),$$
$$\gamma = \begin{cases} (9m^4+1)/2 & ((m,3)=1) \\ (m^4+81)/6 & ((m,3)\neq 1); \end{cases}$$

$$(9) \qquad \frac{125}{24}mb_i^3+\frac{26}{25}m+\gamma \leqq 3^{i-1}C \leqq \frac{125}{18}mb_i^3+\frac{m}{3} \qquad (i=1,\cdots,5),$$

where γ has some value similar in form to those of (8).[9]

An inspection will verify that for $m \geqq 7$,

$$(10) \qquad b_i < 7m \qquad (i=1,2,3)$$

for all the[10] b_i of papers A. Replacing b_i by $7m$ in the right hand side of (8), we get

$$(11) \qquad 3^{i-1}C < 520m^4 \qquad (i=1,2,3).$$

Whence $C < 60m^4$. We also seek a value for 3^{2n} which satisfies[11]

$$\frac{3^{2n}}{b_i(3A_i)^{1/2}} \geqq 8$$

and hence which satisfies

$$(12) \quad 3^{4n} \geqq \left[\frac{384b_i}{m}(3^iC-m)-432b_i^4\right]3^{2n}+\frac{384b_i}{m}(m-6-3b_i)+192b_i^2.$$

---

[8] *Op. cit.* These papers shall henceforth be referred to as papers A.

[9] f and F originally contained terms of the form β_i divided by a power $p \neq p(n)$ of 3, but for $n=\nu$ these terms become negligible.

[10] In his paper, Webber did not list the b_i corresponding to the case $m=18e+12$, i. e. $m=3a$, a even and $\equiv 1 \pmod 3$. They are here supplied by the author: $b_1=20e+11$, $b_2=28e+17$, $b_3=40e+23$.

[11] Dickson, *op. cit.*, p. 7, (28).

For convenience we state the following lemma.

LEMMA 4. *For a positive α, the inequality $x^2 > \alpha x + \beta$ is satisfied by* $x > M = \max(\alpha, \beta) + 1$.

For each of three cases, β negative, $\beta > \alpha$, and $\beta < \alpha$, the proof may be made by substituting M for x.

If we write (12) in the form $x^2 > \alpha x + \beta - \delta$, where β is the sum of all the positive terms free of $3^{2\nu}$, we know that this inequality is satisfied for $x > \max(\alpha, \beta) + 1$, by Lemma 4. But by (10) and (11) we have

$$5{,}000{,}000\, m^4 > 384 b_i 3^i C/m > \max(\alpha, \beta) + 1.$$

Hence we have for $m \geqq 7$,

$$C 3^{3\nu} < 10^{12} m^{10} = C_1.$$

A similar argument for (9) produces a smaller constant than C_1.

THE UNIVERSITY OF CALIFORNIA.
BERKELEY, CALIFORNIA.

ON A PROBLEM OF PÓLYA.

By Norman Levinson.[1]

1. Pólya[2] has set the following problem:

If $f(z)$ is an entire function uniformly bounded at $z = 0, \pm 1, \pm 2, \cdots$, and if

$$\overline{\lim_{r\to\infty}} \frac{\log |f(re^{i\theta})|}{r} \leq 0 \tag{1.0}$$

then $f(z)$ is a constant.

Here we shall prove the following theorem which will yield immediately a generalization of Pólya's results.

Theorem I. *Let $g(z) = g(x+iy)$ be an entire function such that*

$$\overline{\lim_{r\to\infty}} \frac{\log |g(re^{i\theta})|}{r} \leq \pi \tag{1.1}$$

and

$$g(iy) = O(e^{\pi|y|}) \tag{1.2}$$

as $|y| \to \infty$. If

$$g(z_n) = O(1) \tag{1.3}$$

as $|n| \to \infty$, where $\{z_n\}$ is a sequence of complex quantities such that

$$|z_n - n| \leq a, \qquad -\infty < n < \infty \tag{1.4}$$

for some positive integer a, and

$$|z_n - z_m| \geq \delta \tag{1.5}$$

for $n \neq m$ and for some fixed $\delta > 0$, then

$$g(x) = O(|x|^A) \tag{1.6}$$

where A depends only on a.

---

[1] National Research Fellow.

[2] *Jahresbericht der Deutschen Mathematischen Vereinigung*, Bd. 40 (1931), 2te Abteilung, p. 80, problem 105. Solutions have been given by Tschakaloff, by Szegö, and by Pólya. Paley and Wiener in "Fourier transforms in the complex domain," *American Mathematical Society Colloquium Publications*, vol. 19, pp. 81-83, have also given a solution.

As a corollary of this theorem we have the extension of Pólya's result.

THEOREM II. *Let $f(z)$ satisfy condition* (1.0) *and let*

$$f(z_n) = O(1) \tag{1.7}$$

as $|n| \to \infty$, where $\{z_n\}$ is a set of complex quantities such that

$$|z_n - n\alpha| \leq \beta, \qquad -\infty < n < \infty \tag{1.8}$$

for some positive α and β. Then $f(z)$ is a constant.

2. We require several lemmas in proving Theorem I.

LEMMA 1. *Let $\{z_n\}$ satisfy conditions* (1.4) *and* (1.5). *We also assume $|z_n| \geq 1$. If*

$$F(z) = (z - z_0) \prod_1^\infty \left(1 - \frac{z}{z_n}\right)\left(1 - \frac{z}{z_{-n}}\right) \tag{2.0}$$

then

$$|F(z)| < A_1(|z| + 1)^{20a} e^{\pi|y|}, \tag{2.1}$$

$$\left|\frac{F(z)}{z - z_n}\right| < A_2(|z| + 1)^{20a} e^{\pi|y|}, \tag{2.2}$$

$$|F(z)| > \frac{A_3 e^{\pi|y|}}{(|z| + 1)^{20a}} \text{ for } |y| > 2a, \tag{2.3}$$

and

$$|F'(z_n)| > \frac{A_4 \delta^{8a+1}}{(|n| + 1)^{8a+1}}, \qquad -\infty < n < \infty, \tag{2.4}$$

where A_1, A_2, A_3, and A_4 depend only on a.

Since $F(-z)$ satisfies all the requirements we assume with no loss of generality that $x \geq 0$. We begin by proving (2.1) for $|z| \geq 4a$. Clearly if $|n| > 2a$

$$\left|\frac{1 - \dfrac{z}{z_n}}{1 - \dfrac{z}{n}}\right| \leq 1 + \left|\frac{z(z_n - n)}{z_n(n - z)}\right|$$

$$\leq 1 + \frac{a|z|}{|n - z|(|n| - a)} \leq \exp(2a|z| / |n| \; |n - z|).$$

Or if N is defined by $N - \frac{1}{2} \leq |z| < N + \frac{1}{2}$, then

$$\prod_{2a+1}^\infty \left|1 - \frac{z}{z_n}\right| \left|1 - \frac{z}{z_{-n}}\right|$$

$$\leq \prod_{2a+1}^\infty \left|1 - \frac{z^2}{n^2}\right| \left|\frac{1 - \dfrac{z}{z_N}}{1 - \dfrac{z}{N}}\right| \exp[2a|z| \Sigma''(1/|n| \; |n - z|)]$$

where Σ'' indicates the sum from $-\infty$ to ∞ with the terms $n=0$ and $n=N$ omitted. Clearly

$$\begin{aligned}\Sigma'' \frac{1}{|n|\;|n-z|} &\leq \sum_{-\infty}^{-N} \frac{1}{n^2} + \frac{1}{|z|} \sum_{-N}^{-1} \frac{1}{|n|} \\ &+ \frac{3}{|z|} \sum_{1}^{[\frac{1}{2}N]} \frac{1}{n} + \frac{3}{|z|} \sum_{[\frac{1}{2}N]+1}^{N-1} \frac{1}{N-n-\frac{1}{2}} \\ &+ \frac{1}{|z|} \sum_{N+1}^{2N} \frac{1}{(n-N-\frac{1}{2})} + \sum_{2N+1}^{\infty} \frac{1}{(n-N-\frac{1}{2})^2} \\ &\leq \frac{10 \log |z|}{|z|}.\end{aligned}$$

Therefore

$$\prod_{2a+1}^{\infty} \left|1-\frac{z}{z_n}\right| \left|1-\frac{z}{z_{-n}}\right| \leq 2\,|z|^{20a} \left|\frac{z_N-z}{z-N}\right| \left|\frac{\sin \pi z}{\pi z}\right| \prod_{1}^{2a} \left|1-\frac{z^2}{n^2}\right|^{-1}.$$

Or

$$|F(z)| \leq 2\,|z|^{20a} \frac{|z-z_0|\;|z_N-z|\;|\sin \pi z| \prod_{1}^{2a} \left|1-\dfrac{z}{z_n}\right| \left|1-\dfrac{z}{z_{-n}}\right|}{\pi\,|z|\;|N-z| \prod_{1}^{2a} \left|1-\dfrac{z^2}{n^2}\right|}.$$

Recalling that $|z| \geq 4a$ and $|z_n| \geq 1$, we have

$$\begin{aligned}|F(z)| &\leq 10\,|z|^{20a} \frac{(|z|+3a)^{4a+1}(2a)^{4a+1}e^{\pi|y|}}{(|z|-2a)^{4a+1}} \\ &< A\,|z|^{20a}\,e^{\pi|y|}\end{aligned}$$

where A depends only on a. This holds for $|z| \geq 4a$. If $|z| < 4a$ we can extend this result by observing that $F(z)$ being analytic takes its extreme value on the boundary. Thus we get (2.1).

As regards $F(z)/(z-z_n)$ we observe that for $|z-z_n| \geq 1$, (2.2) is a consequence of (2.1). For $|z-z_n| < 1$ we use the fact that $F(z)/(z-z_n)$ takes its extreme value on $|z-z_n| = 1$. This proves (2.2).

The proof of (2.3) is similar to that of (2.1). Here we consider

$$\left|\frac{1-\dfrac{z}{n}}{1-\dfrac{z}{z_n}}\right| = \left|1-\frac{z(z_n-n)}{n(z_n-z)}\right|.$$

If $-\infty < n < |z|-2a$ or if $|z|+2a < n < \infty$, then

$$\left|\frac{1-\dfrac{z}{n}}{1-\dfrac{z}{z_n}}\right| \leq 1+\frac{a\,|z|}{|n|(|z-n|-a)} \leq \exp(2a\,|z|/|n|\;|z-n|).$$

If we now proceed in a manner similar to that used above and remember that $|y| > 2a$ we get

$$\left|\frac{1}{F(z)}\right| \leq \frac{(|z|+1)^{20a}}{4A_3 |\sin \pi z|}$$

which gives (2.3) immediately.

In proving (2.4) we observe that

$$F'(z_n) = -\frac{1}{z_n}(z_n - z_0) \prod_{k=1}^{\infty}{}' \left(1 - \frac{z_n}{z_k}\right)\left(1 - \frac{z_n}{z_{-k}}\right)$$

where the prime on the product indicates that the term zero is omitted. If $n > 3a$,

$$\begin{aligned} |F'(z_n)| &\geq \frac{n-2a}{n+a} \prod_{1}^{n-2a-1} \left(\frac{n-k-2a}{k+a}\right)\left(\frac{n+k-2a}{k+a}\right) \\ &\quad \times \prod_{n+2a+1}^{\infty} \left(\frac{k-n-2a}{k+a}\right)\left(\frac{k+n-2a}{k+a}\right)\left(\frac{\delta}{n+3a}\right)^{8a+1} \\ &\geq \tfrac{1}{4}\left(\frac{\delta}{n+3a}\right)^{8a+1} \prod_{1}^{a} \frac{k^2(n+2a+k)^2}{(n-2a-1+k)^2} \\ &\quad \times \prod_{1}^{n-2a-1} \left|1 - \frac{(n-2a)^2}{k^2}\right| \prod_{n+2a+1}^{\infty} \left(1 - \frac{(n-2a)^2}{k^2}\right). \end{aligned}$$

If we observe that $|\sin \pi(x+k)|/|x|$ approaches π as $x \to 0$ we get, recalling that a is an integer,

$$\prod_{1}^{n-2a-1} \left|1 - \frac{(n-2a)^2}{k^2}\right| \prod_{n+2a+1}^{\infty} \left(1 - \frac{(n-2a)^2}{k^2}\right) = \tfrac{1}{2} \prod_{n-2a+1}^{n+2a} \left|1 - \frac{(n-2a)^2}{k^2}\right|^{-1}.$$

Thus.

$$|F'(z_n)| \geq \tfrac{1}{8}\left(\frac{\delta}{n+3a}\right)^{8a+1} \prod_{n-2a-1}^{n+2a} \left|1 - \frac{(n-2a)^2}{k^2}\right|^{-1} \geq \tfrac{1}{8}\left(\frac{\delta}{n+3a}\right)^{8a+1}.$$

This proves (2.4) for $n > 3a$. Clearly the same method can be used for $0 \leq n \leq 3a$ and since $F(-z)$ is of the same form as $F(z)$ it holds for all values of n.

LEMMA 2. *If $\psi(z)$ is an entire function satisfying* (1.0) *and if $\psi(x)$ is uniformly bounded for real x, then $\psi(z)$ is a constant.*

That $\psi(z)$ is bounded in each half-plane (upper and lower) follows from a well known result of Pragmén and Lindelöf.[3] Thus $\psi(z)$ is a constant.

3. In proving Theorem I we make use of the Pragmén-Lindelöf function,

[3] Pólya-Szegö, *Aufgaben und Lehrsätze*, vol. 1 (1925), p. 147, problem 135.

$h(\theta)$. Since we are dealing exclusively with entire functions of order 1, the following definition will suffice for use here,

$$h(\theta) = \lim_{r\to\infty} \frac{\log |f(re^{i\theta})|}{r}. \tag{3.0}$$

The behaviour of this function is characterized by

THEOREM [4] A. *If* $\theta_1 < \theta_2 < \theta_3$ *and* $\theta_3 - \theta_1 < \pi$ *then*

$$h(\theta_1)\sin(\theta_3-\theta_2) + h(\theta_3)\sin(\theta_2-\theta_1) \geq h(\theta_2)\sin(\theta_3-\theta_1). \tag{3.1}$$

Proof of Theorem I. We can assume that $|z_n| \geq 1$ for if this is not the case then it becomes true on discarding a finite number of z_n. The new set which can again be called $\{z_n\}$ clearly satisfies (1.4) for some new a. Thus

$$F(z) = (z - z_0)\prod_1^\infty \left(1 - \frac{z}{z_n}\right)\left(1 - \frac{z}{z_{-n}}\right)$$

satisfies the requirements of Lemma 1.

We assume that $g(z)$ has an infinite number of zeros (otherwise the theorem is trivial by an application of the Hadamard factorization theorem). If we divide out $8a + 3$ of these zeros it is obvious that we obtain a function $g_1(z)$ such that

$$g_1(z_n) = O(|z_n|^{-8a-3}), \qquad |n| \to \infty.$$

Thus by (2.4)

$$\sum_{-\infty}^{\infty} \frac{g_1(z_n)F(z)}{F'(z_n)(z-z_n)}$$

is absolutely convergent. Moreover

$$H(z) = g_1(z) - \sum_{-\infty}^{\infty} \frac{g_1(z_n)F(z)}{F'(z_n)(z-z_n)}$$

vanishes at all points $z = z_n$ and by (1.1) and (2.2) we have

$$\varlimsup_{r\to\infty} \frac{\log |H(re^{i\theta})|}{r} \leq \pi. \tag{3.2}$$

Thus by the Hadamard factorization theorem

$$H(z) = F(z)\psi(z) \tag{3.3}$$

where $\psi(z)$ is an entire function of at most order 1.

---

[4] Titchmarsh, *Theory of Functions*, Oxford Press (1932), p. 184.

By (1.2) and (2.2) it follows that

$$H(iy) = O(|y|^{20a} e^{\pi|y|}).$$

By (3.3) and (2.3) it follows therefore that

$$\psi(iy) = O(|y|^{40a}). \tag{3.4}$$

We shall now investigate separately the various possible forms of $\psi(z)$.

Case 1. Let us suppose $\psi(z)$ is a polynomial. By (3.4) it is of at most degree $40a$. We recall that

$$g_1(z) = \psi(z)F(z) + \sum_{-\infty}^{\infty} \frac{g_1(z_n)F(z)}{F'(z_n)(z-z_n)}.$$

By (2.1) and (2.2) with $\psi(z)$ a polynomial of degree at most $40a$ it is clear that

$$g_1(x) = O(|x|^{60a}).$$

Therefore

$$g(x) = O(|x|^{68a+3})$$

and the theorem is proved for this case.

Case 2. Let us suppose that $\psi(z)$ is not a polynomial but that

$$\overline{\lim_{|x|\to\infty}} \frac{\log|\psi(x)|}{|x|} \leq 0.$$

Clearly (3.4) gives us

$$\overline{\lim_{|y|\to\infty}} \frac{\log|\psi(iy)|}{|y|} \leq 0. \tag{3.5}$$

Thus applying (3.1) to each quadrant we get

$$h(\theta) = \overline{\lim_{r\to\infty}} \frac{\log|\psi(re^{i\theta})|}{r} \leq 0 \tag{3.6}$$

for all values of θ.

We can now conclude that $\psi(z)$ cannot have a finite number of zeros because if it did then the Hadamard factorization theorem would give

$$\psi(z) = P(z)e^{cz}$$

where $P(z)$ is a polynomial. By (3.6), $c=0$ and therefore $\psi(z)$ is a polynomial contrary to hypothesis.

Thus $\psi(z)$ has an infinite number of zeros and if we divide out $40a$ of these we obtain a function $\psi_1(z)$ such that by (3.4), $\psi_1(iy) = 0(1)$ as

$|y| \to \infty$. Since $\psi_1(z)$ satisfies (3.6), it follows from Lemma 2 that $\psi_1(iz)$ is a constant. But this means $\psi(z)$ is a polynomial which is contrary to our assumption. Thus Case 2 is impossible.

Case 3. Finally we suppose that

$$\overline{\lim_{|x|\to\infty}} \frac{\log|\psi(x)|}{|x|} = \alpha > 0. \tag{3.7}$$

With no loss of generality we can assume that (3.7) holds for $x > 0$ (for otherwise we could deal with $\psi(-x)$). Thus $h(0) = \alpha > 0$. Using (3.5) and applying (3.1) to $0 < \theta < \frac{1}{2}\pi$ and $-\frac{1}{2}\pi < \theta < 0$, we have

$$h(\theta) \leq \alpha \cos\theta, \qquad -\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi. \tag{3.8}$$

If we set $\theta_1 = -\theta$, $\theta_3 = \theta$, and $\theta_2 = 0$ in (3.1) we have

$$h(-\theta)\sin\theta + h(\theta)\sin\theta \geq \alpha\sin 2\theta, \qquad -\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi.$$

Or

$$h(-\theta) + h(\theta) \geq 2\alpha\cos\theta.$$

Combining this with (3.8) we have

$$h(\theta) = \alpha\cos\theta, \qquad -\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi.$$

That is

$$\overline{\lim_{r\to\infty}} \frac{\log|\psi(re^{i\theta})|}{r} = \alpha\cos\theta, \qquad -\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi. \tag{3.9}$$

By (2.3) and (3.2) it follows that α cannot be infinite. Let us take $\theta_0 = \tan^{-1}\pi/\alpha$. Then using (2.1), (2.3), and (3.3) we have

$$\begin{aligned}\overline{\lim_{r\to\infty}} \frac{\log|H(re^{i\theta_0})|}{r} &= \lim_{r\to\infty} \frac{\log|F(re^{i\theta_0})|}{r} + \overline{\lim_{r\to\infty}} \frac{\log|\psi(re^{i\theta_0})|}{r} \\ &= \pi\sin\theta_0 + \alpha\cos\theta_0 = (\pi^2 + \alpha^2)^{\frac{1}{2}}.\end{aligned}$$

But by (3.2), α must be zero, contrary to hypothesis. Thus Case 3 is impossible.

These three cases exhaust all possibilities. Only Case 1 is possible and this leads immediately to the completion of the theorem.[5]

We can now readily prove Theorem II. We observe that there is no restriction in assuming that $\alpha > 3\beta$ in (1.8), for if this is not the case then let N be an integer so large that $N\alpha > 3\beta$. We can then use the sequence

---

[5] Clearly, condition (1.1) need only hold for θ close to $\frac{1}{2}\pi$ and $-\frac{1}{2}\pi$, and $g(z)$ be known to be of order 1, in order that the above proof of Theorem I go through.

$\{z_{nN}\}$ in place of $\{z_n\}$ and $\alpha_1 = \alpha N$ in place of α in (1.8) and now we have $\alpha_1 > 3\beta$.

Thus assuming $\alpha > 3\beta$, we have, if $n \neq m$,

$$\begin{aligned} |z_n - z_m| &\geq |n\alpha - m\alpha| - |z_n - n\alpha| - |z_m - m\alpha| \\ &\geq \alpha - 2\beta > \beta. \end{aligned} \tag{3.10}$$

We observe that $f(z/\alpha)$ is uniformly bounded at the points $\{z_n/\alpha\}$. By (1.8)

$$\left| \frac{\alpha}{z_n} - n \right| \leq \frac{\beta}{\alpha}. \tag{3.11}$$

And by (3.10)

$$\left| \frac{z_n}{\alpha} - \frac{z_m}{\alpha} \right| \geq \frac{\beta}{\alpha}.$$

Thus $f(z/\alpha)$ satisfies the requirements of Theorem I and therefore

$$f\left(\frac{x}{\alpha}\right) = O(|x|^A), \qquad |x| \to \infty,$$

for some A. Or

$$f(x) = O(|x|^A).$$

By the Hadamrd factorization theorem, Theorem II is trivial if $f(z)$ has only a finite number of zeros. If we assume it has an infinite number we divide out A of these zeros and apply Lemma 2, which shows at once that $f(z)$ is an algebraic polynomial and therefore cannot have an infinite number of zeros. This completes the proof of Theorem II.

PRINCETON UNIVERSITY AND THE
INSTITUTE FOR ADVANCED STUDY.

GÉOMÉTRIE DES SYSTÈMES DE CHOSES NORMÉES.

Par V. GLIVENKO.

I. Introduction.

1. On sait que, dans plusieurs questions d'Analyse, il est permis de ne prendre pas en considération les ensembles de mesure nulle, de sorte qu'il y est permis d'identifier les deux ensembles mesurables A et B chaque fois où les points de A n'appartenant pas à B et ceux de B n'appartenant pas à A forment un ensemble de mesure nulle. Autrement dit, on peut y remplacer les ensembles mesurables par leurs *types métriques*.

Les types métriques eux-mêmes n'étant pas des ensembles, il est toutefois naturel d'établir entre eux des relations propres aux ensembles. Ainsi, on peut dire qu'un type a fait partie d'un type b, en signes $a \subset b$, si tout ensemble A du type a est contenu, à un sous-ensemble de mesure nulle près, dans tout ensemble B du type b. Pareillement, on peut introduire la notion de la partie commune ab de deux types a et b, celle de la somme $a + b$ de ces types etc.

Nous exprimons tout cela en disant que les types métriques, de même que les ensembles eux-mêmes, forment un système de choses, dont nous allons préciser la définition.

2. Nous appelons *système de choses* un ensemble S d'éléments $a, b, c, \cdots$ lorsque les conditions suivantes sont remplies:

1°. L'ensemble S contient des couples d'éléments a, b liés entre eux par une relation $a \subset b$ telle que:

$$a \subset b \text{ et } b \subset a \text{ entraîne } a = b \text{ et inversement;}$$
$$a \subset b \text{ et } b \subset c \text{ entraîne } a \subset c.$$

2°. A tout couple d'éléments a, b de l'ensemble S correspond un élément ab de S tel que

$$ab \subset a,$$
$$ab \subset b,$$

et que, x étant un élément de S,

$$x \subset a \text{ et } x \subset b \text{ entraîne } x \subset ab.$$

3°. A tout couple d'éléments a, b de l'ensemble S correspond un élément $a+b$ de S tel que

$$a \subset a+b,$$
$$b \subset a+b,$$

et que, y étant un élément de S,

$$a \subset y \text{ et } b \subset y \text{ entraîne } a+b \subset y.$$

4°. L'ensemble S contient un élément 0 tel que, quel que soit l'élément z de S,

$$0 \subset z.$$

Il est aisé de voir que les propriétés ci-dessus définissent les éléments ab, $a+b$ et 0 d'une manière univoque.

3. A tout ensemble mesurable correspond un nombre bien défini, sa mesure. De même, à tout type métrique a on peut attacher un nombre bien défini $|a|$, mesure d'un quelconque ensemble de ce type.

Il y a cependant une différence essentielle entre ces deux cas. Il existe plusieurs ensembles ayant la mesure nulle, mais il n'y a qu'un seul type $a=0$ tel que $|a|=0$; c'est porquoi tout ensemble peut être augmenté sans changer sa mesure, ce qui est impossible pour les types.

Nous exprimons tout cela en disant que les types métriques forment le système de choses normées.

4. Nous disons que S est le système de choses *normées* lorsqu'à toute chose a de S correspond un nombre non négatif $|a|$, *norme* de cette chose, possédant les propriétés suivantes:

$$a \subset b \text{ et } a \neq b \text{ entraîne } |a| < |b|;$$
$$|a+b| + |ab| = |a| + |b|;$$
$$|0| = 0.$$

5. Outre le système des types métriques, on connaît plusieurs systèmes de choses normées. En voici quelques exemples.

1°. Dans la théorie des probabilités, il est quelquefois favorable d'identifier les deux événements chaque fois où la probabilité que l'un d'eux se produit tandis que l'autre ne se produit pas, est nulle. Autrement dit, il y est favorable de remplacer les événements par leurs *types stocastiques*.

Les types stocastiques forment un système de choses normées si l'on con-

vient d'écrire $a \subset b$ chaque fois où, A étant un événement arbitraite du type a et B un événement arbitraire du type b, la probabilité que A se produit tandis que B ne se produit pas, est nulle. Ici, $|a|$ est la probabilité d'un quelconque événement du type a.

2°. Dans la logique mathématique, il est quelquefois favorable d'identifier les propositions équivalentes. Autrement dit, il y est favorable de remplacer les propositions par leurs *types logiques*.

Les types logiques forment un système de choses normées si l'on convient d'écrire $a \subset b$ chaque fois où, A étant une proposition arbitraire du type a et B une proposition arbitraire du type b, la proposition A implique la proposition B. Ici, $|a|$ est la valeur logique d'une quelconque proposition du type a, égale à *un* pour les proposition vraies et égale à *zéro* pour les propositions fausses.

3°. Prenons les domaines bornés aux frontières quarrables, et considérons ces domaines conjointement avec leurs frontières. Les ensembles fermés ainsi obtenus et l'ensemble vide forment un système de choses normées où $a \subset b$, ab, $a + b$ et 0 ont le sens usuel sauf le cas où la partie commune de a et de b ne contient aucun domaine: alors, on prend pour ab l'ensemble vide. Ici, $|a|$ est l'étendue de a.

4°. Prenons un anneau d'ensembles finis. Ceux-ci et l'ensemble vide forment un système de choses normées où $a \subset b$, ab, $a + b$ et 0 ont le sens usuel. Ici, $|a|$ est le nombre d'éléments de a.

5°. Les nombres entiers positifs forment un système de choses normées où $a \subset b$ signifie que a est un diviseur de b, de sorte que ab est le plus grand diviseur commun de a et de b, $a + b$ est le plus petit multiple commun de ces nombres, 0 est le nombre 1. Ici, on a $|a| = \log a$.

6°. Les nombres non négatifs forment un système de choses normées où $a \subset b$ signifie que a ne dépasse pas b, de sorte que ab est le plus petit des nombres a et b, $a + b$ est le plus grand de ces nombres, 0 est le nombre 0. Ici, on a simplement $|a| = a$.

6. Abordons maintenant la question principale dont nous nous occupons. On verra que tout système de choses normées est un espace métrique. Cela résulte du fait qu'on peut y former une expression, à savoir $|a + b| - |ab|$, qui possède tous les propriétés de la distance de a et de b.

C'est précisement l'étude de la structure de cet espace qui est le but du présent article. Nous définirons un espèce d'espaces métriques que nous appelerons espaces *presque ordonnés*, et nous démontrerons les deux théorèmes suivants:

THÉORÈME DIRECT. *Tout système de choses normées est un espace métrique presque ordonné où la distance de a et de b est égale à*

$$| a + b | - | ab |.$$

THÉORÈME INVERSE. *Tout espace métrique presque ordonné est un système de choses normées où l'expression*

$$| a + b | - | ab |$$

est égale à la distance de point a et de point b.

7. Il nous paraît que telles considérations, où les types métriques, par exemple, se présentent comme des points d'un espace, sont dignes d'intérêt puisque les types métriques peuvent être définis effectivement comme des éléments indépendants, en laissant de côté toute la théorie des ensembles mesurables.

Convenons de désigner, en effet, par $A, B, \cdots$ les sommes finies et bornées d'intervalles, et par $| A |, | B |, \cdots$ les sommes de ses longueurs; ces sommes d'intervalles forment un espace métrique où la distance de A et de B est égale à

$$| A + B | - | AB |.$$

L'espace en question n'est pas *complet*, mais nous pouvons le faire complet en y ajoutant des nouveaux éléments à l'aide des suites convergentes, en se servant d'un procédé bien connu, de la même manière qu'on obtient tous les nombres réels à partir des nombres rationnels.[1] Les éléments de l'espace ainsi complété seront précisement les types métriques.

8. Il est à remarquer, entre autres, que tous les exemples cités plus haut, des systèmes de choses normées, possèdent une propriété importante.

On sait que les propriétés 1°-4° du n°2, caractéristiques pour les systèmes de choses, ont pour conséquence la relation suivante:

$$(1) \qquad ac + bc \subset (a + b)c.$$

Quant à la relation inverse,

$$(2) \qquad (a + b)c \subset ac + bc,$$

elle est, au contraire, indépendante des propriétés en question. Il existe, en effet, des systèmes de choses dans lesquels la relation (2) n'est pas nécessaire-

[1] Cf. F. Hausdorff, *Mengenlehre*, 1927, p. 106.

ment remplie. L'exemple bien connu d'un tel système est celui du système des corps convexes. Si, a et b étant deux corps convexes quelconques, on attribue à $a \subseteq b$, à ab et à 0 le sens usuel et si l'on prend pour $a+b$ le plus petit corps convexe contenant a et b, on voit sans peine que les conditions 1°-4° du n°2 y seront remplies tandis qu'il n'en sera pas, en général, pour la relation (2).

En d'autres termes, convenons de dire qu'un système de choses est *distributif* si, quels que soient a, b, c, on a nécessairement

$$ac + bc = (a+b)c.$$

Ce-ci est équivalent à la couple des relations (1) et (2).

La propriété importante, mentionnée ci-dessus, des exemples cités plus haut, consiste en ce que c'étaient toujours des exemples des systèmes distributifs. On pourrait croire que c'est inévitable pour les systèmes de choses normées. Mais on verra dans la suite qu'il existe aussi des systèmes des choses normées qui ne sont pas distributifs. Cependant, les systèmes distributifs de choses normées méritent, sans doute, une attention particulière et nous leur consacrérons, dans ce qui suit, une étude détaillée.

II. Espaces métriques presque ordonnés.

9. Rappolons qu'un ensemble D est dit *espace métrique*,[2] et que ses éléments $a, b, c, \cdots$ se nomment *points*, lorsqu'à tout couple d'éléments a, b de l'ensemble D correspond un nombre non négatif (a, b), *distance* de point a et de point b, possédant les propriétés suivantes:

$$\begin{aligned}&(a, b) = 0 \text{ entraîne } a = b \text{ et inversement;}\\&(a, b) = (b, a);\\&(a, b) \leq (a, c) + (c, b).\end{aligned}$$

Quand on a $(a, b) = (a, c) + (c, b)$, on dit que le point c se trouve *entre* les points a et b. Il est aisé de voir qu'entre les points a et b se trouvent, en particulier, a et b eux-mêmes.

Nous disons qu'un espace métrique D est *presque ordonné* s'il contient un point que nous appelerons *origine* et qui possède les propriétés suivantes. Convenons de dire qu'un point a est plus *prochain* qu'un point b, ou que b est plus *lointain* que a, si a se trouve entre l'origine et b. Alors:

1°. Quels que soient les deux points a et b, de D, et quel que soit le point c se trouvant entre a et b, chaque point plus prochain que a et b est aussi plus

[2] Il s'agit des espaces découverts par M. M. Fréchet et qu'il a nommé *espaces* (D).

prochain que c; et, de même, chaque point plus lointain que a et b est aussi plus lointain que c.

2°. Quels que soient les deux points a et b, de D, parmi les points se trouvant entre a et b il existe un qui est le plus prochain et il existe un autre qui est le plus lointain.

Pour avoir un exemple d'espace métrique presque ordonné, prenons un ensemble arbitraire de nombres réels, où $(a, b) = | a - b |$. Il est aisé de voir que le rôle de l'origine peut être joué par un nombre quelconque appartenant à cet ensemble.

Pour avoir un exemple d'espace métrique qui n'est pas presque ordonné, prenons un ensemble de nombres complexes, contenant au moins les trois nombres qui ne se trouvent pas sur une droite du plan complexe, d'ailleurs arbitraire, et où, comme auparavant, $(a, b) = | a - b |$. Essayons de prendre un certain nombre c de cet ensemble pour l'origine, et considérons deux nombres a et b tels que c, a et b ne se trouvent pas tous les trois sur une droite. D'après la propriété 2° de l'origine, il y doit exister un nombre d se trouvant entre a et b et tel que, pour tout autre nombre x se trouvant entre a et b, on ait

$$| c - d | + | d - x | = | c - x |.$$

En particulier, comme entre les nombres a et b se trouvent a et b eux-mêmes, on doit y avoir

$$| c - d | + | d - a | = | c - a |$$

et

$$| c - d | + | d - b | = | c - b |.$$

Or, il est manifeste que la première de ces égalités n'est possible que si c, d et a se trouvent sur une droite, de même, la seconde n'est possible que si c, d et b se trouvent sur une droite. Ce-ci contredit à notre supposition que c, a et b ne se trouvent pas sur une droite. On voit ainsi qu'il n'y a aucun nombre possédant les propriétés de l'origine. L'espace n'est pas presque ordonné.

III. Systèmes de choses.

10. Nous commençons par établir une série de propositions auxilières concernant les systèmes de choses et qui nous seront utiles dans notre théorie. La plupart de ces propositions est bien connue dans la Logique mathématique. Les autres, même quand elles ne se rencontrent jamais dans les travaux antérieures, peuvent être reproduites sans peine. Ainsi, ce chapitre peut être omis par le lecteur familier avec les méthodes de la Logique mathématique.

Soit donné donc un système de choses $a, b, c, \cdots$. Alors, on a

(3) $$ab = ba,$$

(4) $$a + b = b + a.$$

Ceci est immédiat.

(5) $$aa = a,$$

(6) $$a + a = a.$$

En effet, d'une part, on a $aa \subset a$. D'autre part, $a \subset aa$, ce qui est une conséquence de $a \subset a$. En comparant, on obtient (5). Pareillement, on établit (6).

Lemme I. *Si l'on a*

$$a \subset c \quad et \quad c \subset d$$

on a aussi

$$ab \subset cd \quad et \quad a + b \subset c + d.$$

Démonstration. On a $ab \subset a$ et $ab \subset b$, donc, lorsque les conditions du lemme, $a \subset c$ et $b \subset d$, sont remplies, on a aussi $ab \subset c$ et $ab \subset d$ d'où $ab \subset cd$. Pareillement, on établit que $a + b \subset c + d$.

Lemme II. *Si l'on a*

$$a \subset b,$$

on a aussi, quel que soit c,

$$ac \subset bc \quad et \quad a + c \subset b + c.$$

Démonstration. Lorsque la condition du lemme, $a \subset b$, est remplie, $ac \subset bc$ est, en vertu du Lemme I, une conséquence de $c \subset c$. Il en est de même pour $a + c \subset b + c$.

On a

(7) $$(ac)(bc) = (ab)c,$$

(8) $$(a + c) + (b + c) = (a + b) + c.$$

En effet, on a, d'une part, $(ac)(bc) \subset ab$, ce qui, en vertu du Lemme I, est une conséquence de $ac \subset a$ et $bc \subset b$. De la même manière, on démontre que $(ac)(bc) \subset c$. Par suite,

$$(ac)(bc) \subset (ab)c.$$

D'autre part, on a $(ab)c \subset ac$, ce qui est, en vertu du Lemme II, und con-

séquence de $ab \subset a$. De la même manière, on démontre que $(ab)c \subset bc$. Par suite

$$(ab)c \subset (ac)(bc).$$

En comparant, on obtient (7). Pareillement, on établit (8).

$$ac + bc \subset c, \tag{9}$$

$$c \subset (a+c)(b+c). \tag{10}$$

En effet, on a $ac \subset c$ et $bc \subset c$, d'où (9). Pareillement, on établit (10).

$$ac + bc \subset a + b, \tag{11}$$

$$ab \subset (a+c)(b+c). \tag{12}$$

En effet, on a $ac \subset a$ et $bc \subset b$, d'où, en vertu du Lemme I, on obtient (11). Pareillement on établit (12).

$$ac + bc \subset (a+b)c, \tag{13}$$

$$ab + c \subset (a+c)(b+c). \tag{14}$$

En effet, (13) est une conséquence de (9) et (11). Pareillement, (14) est une conséquence de (10) et (12).

$$ac + bc \subset ab + c, \tag{15}$$

$$(a+b)c \subset (a+c)(b+c). \tag{16}$$

En effet, (15) est une conséquence de (9) et de $c \subset ab + c$. Pareillement, (16) est une conséquence de $(a+b)c \subset c$ et de (10).

LEMME III. *La relation*

$$a \subset b$$

est équivalente à chacune des relations

$$ab = a \quad \textit{et} \quad a + b = b.$$

DÉMONSTRATION. On a, d'une part,

$$ab \subset a.$$

D'autre part, lorsque $a \subset b$, on a, en vertu du Lemme II,

$$a \subset ab.$$

En comparant, on voit que, lorsque $a \subset b$, on a

$$ab = a.$$

Inversement, lorsque $ab = a$, on obtient $a \subset b$ en vertu de $ab \subset b$. Pareillement, on établit l'équivalence de $a \subset b$ et de $a + b = b$.

On a, quel que soit a,

$$0a = 0, \tag{17}$$

$$0 + a = a. \tag{18}$$

En effet, (17) est, en vertu du Lemme III, une conséquence de $0 \subset a$. Il en est de même pour (18).

$$a(ab) = ab \quad \text{et} \quad a + ab = a, \tag{19}$$

$$b(ab) = ab \quad \text{et} \quad b + ab = b, \tag{20}$$

$$a(a + b) = a \quad \text{et} \quad a + (a + b) = a + b, \tag{21}$$

$$b(a + b) = b \quad \text{et} \quad b + (a + b) = a + b. \tag{22}$$

Pour s'en convaincre, remarquons que la relation $ab \subset a$ est une identité; par suite, les relations (19) qui sont, en vertu du Lemme III, équivalentes à $ab \subset a$, sont, eux-aussi, des identités. Pereillement, en partant de l'identité $ab \subset b$ on obtient (20); en partant de l'identité $a \subset a + b$, on obtient (21); en portant de l'identité $b \subset a + b$, on obtient (22).

Lemme IV. *Si l'on a*

$$a \subset c, \ c \subset b \quad \textit{et} \quad a = b,$$

on a aussi

$$a = c = b.$$

Démonstration. Si les conditions du lemme, $c \subset b$ et $a = b$, sont remplies, on a

$$c \subset a;$$

si, en outre, la condition du lemme

$$a \subset c$$

est aussi remplie, on a, en comparant,

$$a = c.$$

Pareillement, on établit que $c = b$.

LEMME V. *Si l'on a*

$$ab = a + b,$$

on a aussi

$$a = b.$$

DÉMONSTRATION. On a $ab \subset a \subset a + b$ et $ab \subset b \subset a + b$, donc lorsque la condition du lemme, $ab = a + b$, est remplie, on a, en vertu du Lemme IV, $ab = a = a + b$ et $ab = b = a + b$. Il en résulte immédiatement que $a = b$.

LEMME VI. *Si l'on a*

$$x \subset c \subset y,$$
$$a \subset x \subset b,$$
$$a \subset y \subset b,$$

on a aussi

$$a \subset c \subset b.$$

DÉMONSTRATION. Si les conditions du lemme, $x \subset c$ et $a \subset x$, sont remplies, on a $a \subset c$. Pareillement, on établit que $c \subset b$.

On a

$$(ab)c = a(bc), \tag{23}$$

$$(a + b) + c = a + (b + c). \tag{24}$$

En effet, on a, d'une part,

$$(ab)c \subset a$$

ce qui est une conséquence de $(ab)c \subset ab$ et $ab \subset a$. D'autre part, on a

$$(ab)c \subset bc$$

ce qui est, en vertu du Lemme II, une conséquence de $ab \subset a$. Par suite,

$$(ab)c \subset a(bc).$$

De la même manière, on démontre la réciproque,

$$a(bc) \subset (ab)c.$$

En comparant, on obtient (23). Pareillement, on établit (24).

$$ab \subset a + b. \tag{25}$$

En effet, (25) est une conséquence soit de $ab \subset a$ et $a \subset a + b$, soit de $ab \subset b$ et $b \subset a + b$.

LEMME VII. *Si le système est distributif, on a*

$$ab + c = (a + c)(b + c).$$

DÉMONSTRATION. La condition de distributivité étant

$$(a + b)c = ac + bc,$$

on en déduit successivement:

$$(a + c)(b + c) = a(b + c) + c(b + c) = (ab + ac) + (cb + c).$$

Il en résulte, en tenant compte de ce que, d'après (19), $cb + c = c$:

$$(a + c)(b + c) = (ab + ac) + c.$$

De là, en vertu de (24),

$$(a + c)(b + c) = ab + (ac + c).$$

Il en résulte, en tenant compte de ce que, d'après (19), $ac + c = c$:

$$(a + c)(b + c) = ab + c.$$

C'est ce qu'il fallait démontrer.

On appelle parfois *première loi distributive* la relation

$$ac + bc = (a + b)c$$

et *seconde loi distributive* la relation

$$ab + c = (a + c)(b + c).$$

Nous venons de démontrer que la première loi distributive a pour conséquence la seconde; de la même manière, on pourrait aussi démontrer la réciproque.

IV. Systèmes de choses normées.

11. Soit donné maintenant un système de choses normées $a, b, c, \cdots$. Etablissons un principe général concernant ces systèmes, à savoir:

PRINCIPE GÉNÉRAL. *Si l'on sait que* $|a| = |b|$ *et que* $a \subset b$, *on peut affirmer que* $a = b$.

En effet, de $a \subset b$ et $a \neq b$ il résulterait $|a| < |b|$.

12. Ce-ci posé, nous pouvons aborder la démonstration des théorèmes fondamentaux mentionnés dans l'Introduction. Dans la suite, ce seront les Théorèmes I et IV.

THÉORÈME I. *Tout système de choses normées est un espace métrique presque ordonné où la distance de a et de b est égale à*

$$|a+b|-|ab|.$$

PREMIÈRE PARTIE DE DÉMONSTRATION. Commençons par démontrer que tout système de choses normées est un espace métrique où la distance de a et de b est égale à $|a+b|-|ab|$, ça veut dire que la distance définie par l'égalité

$$(a, b)=|a+b|-|ab|$$

possède effectivement toutes les propriétés de la distance de deux points d'un espace métrique.

Etablissons d'abord que $(a, b)=0$ entraîne $a=b$ et inversement.

Soit $(a, b)=0$. Ce-ci n'est autre chose que $|ab|=|a+b|$. Or, d'après (25), on a $ab \subset a+b$. En vertu du principe général, on en conclut que $ab=a+b$. Donc, d'après le Lemme V, on a $a=b$. Inversement, soit $a=b$. Autrement dit, considérons la distance (a, a). Elle est égale à $|a+a|-|aa|$. Or, d'après (5) et (6), on a $aa=a$ et $a+a=a$. Donc, on a $(a, a)=0$.

Etablissons maintenant que

$$(a, b)=(b, a).$$

(a, b) est égale à $|a+b|-|ab|$ et (b, a) est égale à $|b+a|-|ba|$. Or, d'après (3) et (4), on a $ab=ba$ et $a+b=b+a$. Donc, on a $(a, b)=(b, a)$.

Il nous reste à établir que

$$(a, b) \leq (a, c)+(c, b).$$

D'après (8), on a $(a+b)+c=(a+c)+(b+c)$ et, d'après (16), $(a+b)c \subset (a+c)(b+c)$. Donc, en tenant compte des propriétés de la norme,

$$\begin{aligned} (26) \qquad |a+b|+|c| &= |(a+b)+c|+|(a+b)c| \\ &\leq |(a+c)+(b+c)|+|(a+c)(b+c)|=|a+c|+|b+c|. \end{aligned}$$

Puis, d'après (7), on a $(ab)c=(ac)(bc)$ et, d'après (15), $ac+bc \subset ab+c$. Donc, en tenant compte des mêmes propriétés de la norme,

$$(27)\qquad \begin{aligned} &|ab| + |c| = |ab + c| + |(ab)c| \\ &\geq |ac + bc| + |(ac)(bc)| = |ac| + |bc|. \end{aligned}$$

En comparant les inégalités (26) ét (27), on obtient

$$|a + b| - |ab| \leq |a + c| - |ac| + |b + c| - |bc|.$$

Or, ce n'est autre chose que $(a, b) \leq (a, c) + (c, b)$.

13. Avant de terminer la démonstration du Théorème I, nous démontrerons les deux autres théorèmes.

Théorème II. *Dans l'espace métrique formé par un système de choses normées, un point c se trouve entre deux points a et b si et seulement si l'on a*

$$ac + bc = c = (a + c)(b + c).$$

Démonstration. Supposons d'abord que c se trouve entre a et b et proposons-nous d'établir la relation

$$ac + bc = c.$$

A cet effet, remarquons qu'on a, d'après notre supposition,

$$(a, b) = (a, c) + (c, b),$$

ou bien que

$$(28)\qquad |a + b| - |ab| = |a + c| - |ac| + |b + c| - |bc|.$$

En vertu de la propriété fondamentale de la norme, on peut remplacer, dans (28), $|a + b|$ par $|a| + |b| - |ab|$, puis $|a + c|$ par $|a| + |c| - |ac|$ et enfin $|b + c|$ par $|b| + |c| - |bc|$, de sorte que (28) s'ecrira comme il suit:

$$(29)\qquad -|ab| = |c| - |ac| - |bc|.$$

Or, d'après (7), on a $(ac)(bc) = (ab)c$. Donc, en vertu de la même propriété de la norme,

$$(30)\qquad -|(ab)c| = |ac + bc| - |ac| - |bc|.$$

Puis, on a $(ab)c \subset ab$, d'où $|(ab)c| \leq |ab|$. Il en résulte d'après les égalités (29) et (30), que

$$(31)\qquad |c| \leq |ac + bc|.$$

D'autre part, d'après (9), on a $ac + bc \subset c$, d'où

$$(32)\qquad |ac + bc| \leq |c|.$$

En comparant les inégalités (31) et (32), on obtient l'égalité

$$|ac + bc| = |c|.$$

Cette dernière égalité et la relation $ac + bc \subset c$ fournissent, en vertu du principe général, la relation à démontrer $ac + bc = c$.

Pareillement, on établit la relation $c = (a + c)(b + c)$.

Inversement, supposons qu'on a

$$ac + bc = c = (a + c)(b + c)$$

et proposons-nous d'établir que c se trouve entre a et b, c'est-à dire que $(a, b) = (a, c) + (c, b)$. A cet effet, démontrons d'abord que notre supposition entraîne les deux relations suivantes:

$$ab = (ac)(bc), \tag{33}$$

$$a + b = (a + c) + (b + c). \tag{34}$$

Quant à la première, on a, d'après (7),

$$(ac)(bc) = (ab)c.$$

Puis, d'après (12), on a $ab \subset (a + c)(b + c)$, d'où, en tenant compte de notre supposition, $ab \subset c$. En vertu du Lemme III, ce-ci équivant à

$$(ab)c = ab.$$

En comparant, on obtient (33). Pareillement, on établit (34).

Cela posé, on a

$$|a + b| - |ab| = |(a + c) + (b + c)| - |(ac)(bc)|.$$

On en déduit, en se servant encore une fois de notre supposition:

$$|a + b| - |ab| = |(a + c) + (b + c)|$$
$$+ |(a + c)(b + c)| - |ac + bc| - |(ac)(bc)|.$$

Ce-ci fournit, en vertu de la propriété fondamentale de la norme:

$$|a + b| - |ab| = |a + c| + |b + c| - |ac| - |bc|,$$

ou, ce qui revient au même,

$$|a + b| - |ab| = |a + c| - |ac| + |b + c| - |bc|.$$

Or, c'est précisemment l'égalité $(a, b) = (a, c) + (c, b)$, ce qui termine la démonstration de notre théorème.

THÉORÈME III. *Lorsque, dans l'espace métrique formé par un système de choses normées, on prend* 0 *pour l'origine, un point a est plus prochain qu'un point b si et seulement si l'on a*

$$a \subset b.$$

DÉMONSTRATION. "a est plus prochain que b" n'est autre chose que "a se trouve entre l'origine et b." En vertu du Théorème II, ce-ci équivant à la couple des relation que voici:

$$0a + ba = a, \tag{35}$$

$$a = (0 + a)(b + a). \tag{36}$$

Or, d'après (17), on a $0a = 0$ et, d'après (18), $0 + ba = ba$, de sorte que (35) peut s'écrire $ba = a$. En vertu du Lemme III, ceci équivaut à $a \subset b$. Quant à (36), elle ne fournit aucune restriction, car ce n'est qu'une identité. Pour s'en convaincre, remarquons que, d'après (18), on a $0 + a = a$ de sorte que (36) peut s'écrire $a = a(b + a)$. Or, en vertu du Lemme III, ce-ci équivant à $a \subset b + a$, ce qui est une identité. Ainsi, la couple des relations (35) et (36) équivaut, elle-aussi, à $a \subset b$.

14. Nous pouvons maintenant terminer la démonstration du Théorème I en établissant que tout système de choses normées est un espace métrique *presque ordonné.*

SECONDE PARTIE DE LA DÉMONSTRATION DU THÉORÈME I. Dans l'espace métrique formé par un système de choses normées, prenons 0 pour l'origine et démontrons que les deux axiomes caractérisant l'espace métrique presque ordonné seront vrais pour cet espace.

1°. Quels que soient les deux points a et b et quel que soit le point c se trouvant entre a et b, chaque point plus prochain que a et b est aussi plus prochain que c et chaque point plus lointain que a et b est aussi plus lointain que c.

En effet, soit x un point plus prochain que a et b, c'est-à dire, d'après le Théorème III, que $x \subset a$ et $x \subset b$. Alors, $x \subset a + c$ et $x \subset b + c$, d'où

$$x \subset (a + c)(b + c).$$

Lorsque c se trouve entre a et b et, par suite,

$$(a + c)(b + c) = c,$$

on en conclut que $x \subset c$, c'est-à-dire que x est plus prochain que c. De même, soit y un point plus lointain que a et b, c'est-à-dire, d'après le Théorème III, que $a \subset y$ et $b \subset y$. Alors, $ac \subset y$ et $bc \subset y$, d'où

$$ac + bc \subset y.$$

Lorsque c se trouve entre a et b et, par suite,

$$c = ac + bc,$$

on en conclut que $c \subset y$, c'est-à-dire que y est plus lointain que c.

2°. Quels que soient les deux points a et b, il existe parmi les points que se trouvent entre a et b un point qui est le plus prochain et il en existe un autre qui est le plus lointain.

En effet, ces points sont ab (le plus prochain) et $a + b$ (le plus lointain).

Premièrement, ils se trouvent effectivement entre a et b. Pour s'en convaincre, remarquons que, d'une part, d'après (19) et (20), on a

$$\begin{aligned} a(ab) &= ab, \\ b(ab) &= ab, \\ a &= a + ab, \\ b &= b + ab, \end{aligned}$$

d'où l'on déduit tout de suite, en se servant de la formule (6) $a + a = a$, que

$$a(ab) + b(ab) = ab = (a + ab)(b + ab).$$

Ceci nous montre que ab se trouve entre a et b. D'autre part, d'après (21) et (22), on a

$$\begin{aligned} a(a + b) &= a, \\ b(a + b) &= b, \\ a + b &= a + (a + b), \\ a + b &= b + (a + b), \end{aligned}$$

d'où l'on déduit tout de suite, en se servant de la formule (5) $a = aa$, que

$$a(a + b) + b(a + b) = a + b = (a + (a + b))(b + (a + b)).$$

Ceci nous montre que $a + b$ se trouve, lui-aussi, entre a et b.

Deuxièmement, si un point c se trouve entre a et b, ab est plus prochain que c et $a + b$ est plus lointain que c. Autrement dit, en tenant compte du Théorème III,

$$ab \subset c \subset a + b.$$

Pour s'en convaincre, remarquons que, d'après (13) et (14), on a

$$(37) \qquad ac + bc \subset (a+b)c \subset c \subset ab + c \subset (a+c)(b+c).$$

Or, lorsque c se trouve entre a et b, on a

$$ac + bc = c = (a+c)(b+c).$$

Il en résulte, en vertu du Lemme IV, que les relations (37) prennent la forme:

$$(38) \qquad ac + bc = (a+b)c = c = ab + c = (a+c)(b+c).$$

Mais la trosième et à la deuxième des relations (38) :

$$c = ab + c,$$
$$(a+b)c = c$$

sont en vertu du Lemme III équivalentes à ce que nous avons à démontrer:

$$ab \subset c \subset a + b.$$

Le Théorème I est ainsi complétement établi.

15. Abordons la démonstration du théorème inverse.

THÉORÈME IV. *Tout espace métrique presque ordonné est un système de choses normées où l'expression*

$$|a+b| - |ab|$$

est égale à la distance du point a et du point b.

PREMIÈRE PARTIE DE LA DÉMONSTRATION. Ecrivons, par définition, $a \subset b$ si le point a est plus prochain que le point b, prenons pour ab le plus prochain des points se trouvant entre a et b, pour $a+b$ le plus lointain de ces points et pour 0 l'origine de l'espace, et posons, enfin, $|a| = (0, a)$.

Démontrons que tous les axiomes caractérisant les systèmes de choses y seront remplis.

1°. La relation $a \subset b$ possède les propriétés suivantes:

$a \subset b$ et $b \subset a$ entraîne $a = b$ et inversement:

$a \subset b$ et $b \subset c$ entraîne $a \subset c$.

Démontrons la première de ces propriétés. Les relations $a \subset b$ et $b \subset a$ ne sont autres choses que

$$(0, a) + (a, b) = (0, b),$$
$$(0, b) + (b, a) = (0, a).$$

En additionnant ces égalités et en remarquant qu'on a $(a, b) = (b, a)$, on obtient $(a, b) = 0$, c'est-à-dire que $a = b$. Inversement, si $a = b$, on a $(a, b) = (b, a)$ et $(0, a) = (0, b)$, de sorte que

$$(0, a) + (a, b) = (0, b),$$
$$(0, b) + (b, a) = (0, a),$$

c'est-à-dire que $a \subset b$ et $b \subset a$.

Démontrons la deuxième propriété. Les relations $a \subset b$, $b \subset c$ et $a \subset c$ sont équivalentes respectivement aux inégalités:

$$(0, a) + (a, b) \leq (0, b),$$
$$(0, b) + (b, c) \leq (0, c),$$
$$(0, a) + (a, c) \leq (0, c).$$

Admettant les deux premières inégalités, on en déduit la dernière, car

$$(0, a) + (a, c) \leq (0, a) + (a, b) + (b, c) \leq (0, b) + (b, c) \leq (0, c),$$

2°. Le point ab possède les propriétés suivantes:

$$ab \subset a,$$
$$ab \subset b,$$
$$x \subset a \text{ et } x \subset b \text{ entraîne } x \subset ab.$$

La dernière de ces propriétés est évidente, puisque le point ab est, par définition, un des points se trouvant entre a et b. Or, dans l'espace métrique presque ordonné, chaque point x qui est plus prochain que a et b est aussi plus prochain que tout point se trouvant entre a et b. En particulier, x est plus prochain que ab.

Pour démontrer les deux premières propriétés, rappelons que a se trouve toujours entre a et b et que b se trouve aussi entre a et b. Or, ab est, par définition, le point qui est plus prochain que tout autre point se trouvant entre a et b. En particulier, ab est plus prochain que les points a et b eux-mêmes.

3°. Le point $a + b$ possède les propriétés suivantes:

$$a \subset a + b,$$
$$b \subset a + b,$$
$$a \subset y \text{ et } b \subset y \text{ entraîne } a + b \subset y.$$

La démonstration est tout-à-fait analogue à la précédente.

4°. Le point 0 possède la propriété que, quel que soit le point z, on a

$$0 \subset z.$$

En effet, cette propriété n'est autre chose que l'égalité évidente:

$$(0, 0) + (0, z) = (0, z).$$

16. Il nous reste à démontrer qu'avec nos conventions tous les axiomes caractérisant la norme sont remplis et qu'on a, de plus,

$$(a, b) = |a + b| - |ab|.$$

Seconde partie de la démonstration du Théorème IV. Le fait que $a \subset b$ et $a \neq b$ entraîne $|a| < |b|$, est la conséquence immédiate du fait que

$$(0, a) + (a, b) = (0, b)$$

et $(a, b) > 0$ entraîne

$$(0, a) < (0, b).$$

Démontrons maintenant que

$$|a + b| + |ab| = |a| + |b|.$$

A cet effet, exprimons $|ab|$ à l'aide de $|a|$, de $|b|$ et de (a, b). On a $ab \subset a$ et $ab \subset b$, c'est-à-dire que

$$(0, ab) + (ab, a) = (0, a),$$
$$(0, ab) + (ab, b) = (0, b).$$

En additionnant, on obtient

$$2(0, ab) + (ab, a) + (ab, b) = (0, a) + (0, b).$$

Or, ab se trouve entre a et b, donc

$$(ab, a) + (ab, b) = (a, b).$$

Par suite,

$$2(0, ab) + (a, b) = (0, a) + (0, b),$$

ou, ce qui revient au même,

$$|ab| = \frac{|a| + |b| - (a, b)}{2}. \tag{39}$$

Pareillement, on établit que

$$|a + b| = \frac{|a| + |b| + (a, b)}{2}. \tag{40}$$

Remarquons en passant que ce sont les généralisations des expressions connues pour *le plus petit des nombres a et b*,

$$\frac{|a| + |b| - |a - b|}{2},$$

et pour *le plus grand de ces nombres,*

$$\frac{|a|+|b|+|a-b|}{2}.$$

En additionnant les égalités (39) et (40), on obtient ce que nous avions à démontrer, à savoir $|a+b|+|ab|=|a|+|b|$.

Enfin, le fait que $|0|=0$ est la conséquence immédiate de ce que $(0,0)=0$.

Quant à l'égalité

$$(a,b)=|a+b|-|ab|,$$

elle est, elle-aussi, une conséquence de (39) et de (40).

Le Théorème IV est ainsi complétement établi.

V. Espaces transitifs.

17. Ayant en vue l'étude particulière des systèmes distributifs de choses normées, nous allons introduire une définition nouvelle. Il existe un espèce d'espaces métriques presque ordonnés qui jouent dans la théorie des systèmes distributifs le méme rôle que les espaces métriques presque ordonnés quelconques jouent dans la théorie des systèmes arbitraires de choses normées. Ce sont les espaces métriques presque ordonnés D que nous appelerons *transitifs* et qui possèdent, par définition, la propriété suivante:

T. *Si un point c, de D, se trouve entre x et y et si tous les deux points x et y se trouvent entre a et b, le point c se trouve, lui-aussi, entre a et b.*

Pour avoir un exemple d'espace métrique presque ordonné transitif, il suffit de rappeler l'exemple déjà cité d'un ensemble arbitraire de nombres réels, où $(x,y)=|x-y|$.

Pour avoir un exemple d'espace métrique presque ordonné qui n'est pas transitif, prenons l'espace formé de cinq points abstraits 0, a, b, c et d, où la distance (x,y) est définie par le Tableau I.

I. (x,y)

$x \backslash y$	0	a	b	c	d
0	0	1	1	1	2
a	1	0	2	2	1
b	1	2	0	2	1
c	1	2	2	0	1
d	2	1	1	1	0

Ici, on a

$$(0, c) + (c, d) = (0, d),$$
$$(a, 0) + (0, b) = (a, b),$$
$$(a, d) + (d, b) = (a, b),$$

de sorte que c se trouve entre 0 et d, que 0 se trouve entre a et b et que d se trouve entre a et b. Cependant, on a

$$(a, c) + (c, b) \neq (a, b),$$

de sorte que c ne se trouve pas entre a et b.

Le fait que cet espace est métrique peut être prouvé par une simple comparaison des distances du Tableau I. Le fait que cet espace est, de plus, presque ordonné s'éclaircira de lui-même dans le Chapitre VII.

VI. Systèmes distributifs.

18. Dans ce qui suit, chaque fois que l'on considérera un espace métrique presque ordonné, les notations du Chapitre IV seront employées, c'est à dire que nous écrirons $a \subset b$ pour exprimer que le point a est plus prochain que le point b, et nous désignerons par ab le plus prochain des points se trouvant entre a et b par $a + b$ le plus lointain de ces points, par 0 l'origine de l'espace et par $|a|$ la distance $(0, a)$. Cette convention faite, nous aurons un théorème qui nous sera utile dans tout ce qui suit.

Théorème V. *La condition nécessaire et suffisante pour qu'un espace métrique presque ordonné D soit transitif est qu'il possède la propriété suivante:*

T bis. *Un point c, de D, se trouve entre deux points a et b si et seulement si l'on a*

$$ab \subset c \subset a + b.$$

Démonstration. Remarquons d'abord que la condition T bis peut s'exprimer plus simplement en y omettant les mots "et seulement si." En effet, nous savons déjà que, lorsque c se trouve entre a et b, on a nécessairement $ab \subset c \subset a + b$.

Supposons maintenant que l'espace en question est transitif, c'est-à-dire que la condition T est remplie, et soit $ab \subset c \subset a + b$. Nous allons voir que c se trouve alors entre ab et $a + b$. En effet, on a

$$(ab, a + b) = |(a + b) + ab| - |(a + b)(ab)|,$$
$$(ab, c) = |c + ab| - |c(ab)|,$$
$$(c, a + b) = |(a + b) + c| - |(a + b)c|.$$

Mais on a, d'après (25), $ab \subset a + b$, et, d'après notre supposition, $ab \subset c$ et $c \subset a + b$. En vertu du Lemme III, on en conclut que

$$\begin{aligned} (a+b)+ab &= a+b \quad \text{et} \quad (a+b)(ab) = ab, \\ c+ab &= c \quad \text{et} \quad c(ab) = ab, \\ (a+b)+c &= a+b \quad \text{et} \quad (a+b)c = c. \end{aligned}$$

Par suite, on a

$$\begin{aligned} (ab, a+b) &= |\, a+b \,| - |\, ab \,|, \\ (ab, c) &= |\, c \,| - |\, ab \,|, \\ (c, a+b) &= |\, a+b \,| - |\, c \,|, \end{aligned}$$

d'où l'on tire ce qu'il fallait déduire:

$$(ab, a+b) = (ab, c) + (c, a+b).$$

Mais ab et $a+b$ se trouvent entre a et b. Par suite, d'après T, le point c se trouve, lui-aussi, entre a et b. On voit ainsi que, si l'on a $ab \subset c \subset a + b$, le point c se trouve entre a et b, ça veut dire que, d'après la remarque ci-dessus, la condition T bis est remplie.

Inversement, supposons que la condition T bis soit remplie, que c se trouve entre x et y et que x et y se trouvent entre a et b. Alors, nous savons qu'on a

$$\begin{aligned} xy &\subset c \subset x + y, \\ ab &\subset x \subset a + b, \\ ab &\subset y \subset a + b. \end{aligned}$$

Ceci peut s'écrire aussi, en vertu du Lemme I et d'après les formules (5) $aa = a$ et (6) $a + a = a$,

$$\begin{aligned} xy &\subset \quad c \quad \subset x + y, \\ ab &\subset \quad xy \quad \subset a + b, \\ ab &\subset x + y \subset a + b. \end{aligned}$$

On en déduit, en vertu du Lemme VI,

$$ab \subset c \subset a + b.$$

Par suite, d'après T bis, c se trouve entre a et b. On voit ainsi que, si le point c se trouve entre x et y, et que x et y se trouvent entre a et b, le point c se trouve, lui-aussi, entre a et b. C'est la condition T, ça veut dire que l'espace est transitif.

Cela posé, on peut donner une propriété caractéristique des systèmes distributifs.

THÉORÈME VI. *La condition nécessaire et suffisante pour qu'un système de choses normées soit distributif est que l'espace métrique presque ordonné formé par ce système soit transitif.*

DÉMONSTRATION. Supposons d'abord que le système en question soit distributif. Soit, dans l'espace formé par ce système,

$$ab \subset c \subset a + b.$$

Alors, en vertu du Lemme III, on a

$$(a + b)c = c = ab + c.$$

Or, le système étant supposé distributif et en tenant compte du Lemme VII, on a

$$(a + b)c = ac + bc,$$
$$ab + c = (a + c)(b + c).$$

Donc, on a

$$ac + bc = c = (a + c)(b + c).$$

Ceci nous montre que c se trouve entre a et b.

Ainsi, on voit que, si l'on a $ab \subset c \subset a + b$, le point c se trouve entre a et b. La réciproque étant toujours vraie, le théorème V nous apprend que l'espace en question est transitif.

Inversement, supposons que, dans un espace métrique presque ordonné, la condition de distributivité,

$$ac + bc = (a + b)c,$$

n'est pas remplie pour au moins trois points a, b, c et démontrons qu'il existe alors un point x tel que $ab \subset x \subset a + b$ et qui cependant ne se trouve pas entre a et b. En vertu du Théorème V, cela suffira pour affirmer que l'espace n'est pas transitif.

Ainsi, soit

$$ac + bc \neq (a + b)c.$$

On sait toutefois, d'après (13), que

$$ac + bc \subset (a + b)c.$$

Tout espace métrique presque ordonné étant un système de choses normées, il en résulte que

$$| ac + bc | < | (a + b)c |.$$

Or, en vertu de la propriété fondamentale de la norme et d'après la formule (7) $(ac)(bc) = (ab)c$, l'expression $|ac + bc|$ peut être remplacée par $|ac| + |bc| - |(ab)c|$, d'où

$$|ac| + |bc| < |(a+b)c| + |(ab)c|.$$

Comme on a $ac \subset a$ et $a \subset a + b$, puis $bc \subset b$ et $b \subset a + b$, et, enfin, $(ab)c \subset ab$ et $ab \subset a + b$, (la dernière des ces relations étant la formule (25)), on a aussi $ac \subset a + b$, $bc \subset a + b$ et $(ab)c \subset a + b$. Autrement dit, en tenant compte du Lemme III, on a $ac = (a+b)(ac)$, $bc = (a+b)(bc)$ et $(ab)c = (a+b)(ab)c$. Donc, on a

$$|(a+b)(ac)| + |(a+b)(bc)| < |(a+b)c| + |(a+b)(ab)c|.$$

D'après les formules (3) $ab = ba$ et (23) $(ab)c = a(bc)$, ceci peut s'écrire encore comme il suit:

$$|((a+b)c)a| + |((a+b)c)b| < |(a+b)c| + |((a+b)c)(ab)|.$$

En vertu de la propriété fondamentale de la norme, les expressions $|((a+b)c)a|$, $|((a+b)c)b|$ et $|((a+b)c)(ab)|$ peuvent être remplacées respectivement par

$$\begin{gathered} |(a+b)c| + |a| - |(a+b)c + a|, \\ |(a+b)c| + |b| - |(a+b)c + b|, \\ |(a+b)c| + |ab| - |(a+b)c + ab|. \end{gathered}$$

Donc, on a

$$|a| - |(a+b)c + a| + |b| - |(a+b)c + b| < |ab| - |(a+b)c + ab|.$$

Comme on a, d'une part, $ab \subset a$ et $a \subset (a+b)c + a$ et, d'autre part, $ab \subset b$ et $b \subset (a+b)c + b$, on a aussi

$$\begin{gathered} ab \subset (a+b)c + a, \\ ab \subset (a+b)c + b. \end{gathered}$$

Autrement dit, en tenant compte du Lemme III,

$$\begin{gathered} (a+b)c + a = ((a+b)c + a) + ab, \\ (a+b)c + b = ((a+b)c + b) + ab. \end{gathered}$$

Donc, on a

$$\begin{aligned} |a| - |((a+b)c + a) + ab| + |b| & \\ - |((a+b)c + b) + ab| &< |ab| - |(a+b)c + ab|. \end{aligned}$$

D'après les formules (4) $a + b = b + a$ et (24) $(a + b) + c = a + (b + c)$, ceci peut s'écrire encore comme il suit:

$$\begin{aligned} &| a | - | ((a + b)c + ab) + a | + | b | \\ &\qquad - | ((a + b)c + ab) + b | < | ab | - | (a + b)c + ab |, \end{aligned}$$

ou, ce qui revient au même,

$$\begin{aligned} &| a | + | b | - 2 | ab | < 2 | ((a + b)c + ab) + a | - | (a + b)c + ab | - | a | \\ &\qquad + 2 | ((a + b)c + ab) + b | - | (a + b)c + ab | - | b |. \end{aligned}$$

En s'appuyant encore une fois sur la propriété fondamentale de la norme, on en obtient:

$$\begin{aligned} &| a + b | - | ab | < | ((a + b)c + ab) + a | - | ((a + b)c + ab)a | \\ &\qquad + | ((a + b)c + ab) + b | - | ((a + b)c + ab)b |. \end{aligned}$$

Ceci n'est autre chose que

$$(a, b) < (a, (a + b)c + ab) + ((a + b)c + ab, b).$$

On voit donc que, si l'on pose

$$x = (a + b)c + ab,$$

on aura

$$(a, b) < (a, x) + (x, b),$$

c'est-à-dire que x ne se trouve pas entre a et b.

Cependant, en s'appuyant sur la relation (25) $ab \subset a + b$, on verra tout de suite que

$$ab \subset x \subset a + b.$$

VII. Exemple de système non distributif de choses normées.

19. Pour avoir un exemple de système non distributif de choses normées, prenons le système des choses abstraites 0, a, b, c, d, où, par définition, toutes ces cinq choses sont différentes et où l'on a

$$0 \subset a, \quad 0 \subset b, \quad 0 \subset c, \quad 0 \subset d,$$
$$a \subset d, \quad b \subset d, \quad c \subset d.$$

Définissons ensuite xy et $x + y$ par les Tableaux II et III.

II xy

$x \backslash y$	0	a	b	c	d
0	0	0	0	0	0
a	0	a	0	0	a
b	0	0	b	0	b
c	0	0	0	c	c
d	0	a	b	c	d

III $x+y$

$x \backslash y$	0	a	b	c	d
0	0	a	b	c	d
a	a	a	d	d	d
b	b	d	b	d	d
c	c	d	d	c	d
d	d	d	d	d	d

Posons aussi

$$|0|=0, \quad |a|=1, \quad |b|=1, \quad |c|=1, \quad |d|=2.$$

Il est aisé de prouver que ce système-ci est effectivement un système de choses. Si $x \subset y$ et $x \neq y$, on a nécessairement $|x| < |y|$, car on a

$$|0|<|a|, \quad |0|<|b|, \quad |0|<|c|, \quad |0|<|d|,$$
$$|a|<|d|, \quad |b|<|d|, \quad |c|<|d|.$$

La propriété fondamentale de la norme

$$|x+y|+|xy|=|x|+|y|$$

se démontre par un simple calcul dont les résultats sont exposés dans le Tableau IV.

IV $|x+y|+|xy|=|x|+|y|$

$x \backslash y$	0	a	b	c	d
0	0	1	1	1	2
a	1	2	2	2	3
b	1	2	2	2	3
c	1	2	2	2	3
d	2	3	3	3	4

Enfin, nous avons déjà vu que, par la définition même, $|0|=0$. Le système en question est donc un système de choses normées.

Ce système n'est pas distributif, car on a

$$(a+b)c = dc = c,$$

tandis que

$$ac+bc = 0+0 = 0.$$

En considérant ce système comme espace métrique, on obtient l'espace que nous avons défini, par le Tableau I, dans le Chapitre V. (En vertu du Théorème I, celui-ci est presque ordonné.)

On peut indiquer une interprétation élégante du système en question. A cet effet, prenons pour 0 l'ensemble vide, pour a, b et c trois points d'une droite, et pour d la droite même. Attribuons à $x \subset y$ et à xy le sens usuel, et prenons pour $x + y$ la plus petite des cinq choses nommées embrassant à la fois x et y. Pour $|x|$, nous prenons le nombre de dimensions de x augmenté de 1.

Un exemple analogue nous fournirait le système de tous les sous-espaces linéaires d'un espace projectif donné.

VIII. Espaces métriques presque ordonnés qui sont simplement ordonnés.

20. En introduisant le nom de l'espace métrique *presque ordonné* nous avions en vue, bien entendu, la notion d'ordre $a \prec b$ dont l'expression précise se compose de deux affirmations, savoir $a \subset b$ et $a \neq b$. Il n'est pas dépourvu d'intérêt de se demander quand l'espace métrique presque ordonné est un espace simplement ordonné avec la même notion d'ordre. Autrement dit, quand l'espace en question est tel que, quels que soient ses points a et b, il existe nécessairement entre eux l'une des trois relations $a \prec b$, $b \prec a$ ou $a = b$, ou bien, ce qui revient au même, l'une des deux relations $a \subset b$ ou $b \subset a$. La réponse nous est donné dans le théorème:

THÉORÈME VII. *La condition nécessaire et suffisante pour qu'un espace métrique presque ordonné D soit simplement ordonné est qu'il possède la propriété suivante:*

O. *Un point c, de D, se trouve entre deux points a et b si et seulement si l'on a, ou bien*

$$a \subset c \subset b,$$

ou bien

$$b \subset c \subset a.$$

DÉMONSTRATION. Rappelons que si, dans un espace métrique presque ordonné, on a $a \subset b$, on a aussi, en vertu du Lemme III, $ab = a$ et $a + b = b$.

Supposons maintenant que l'espace en question est ordonné et soit, pour fixer les idées, $a \subset b$ (ce qui correspond à la première des éventualités de la condition O). Soit c un point se trouvant entre a et b. Alors

$$ab \subset c \subset a + b,$$

et par suite, d'après la remarque précédente

$$a \subset c \subset b.$$

Pareillement, on démontre que si l'on a $b \subset a$, on a aussi $b \subset c \subset a$.

Inversement, soit, pour tout point c se trouvant entre a et b,

$$a \subset c \subset b.$$

Alors, les points a et b se trouvant eux-mêmes entre a et b, on a, en particulier,

$$a \subset b.$$

Pareillement, on démontre que si l'on a $b \subset c \subset a$, on a aussi $b \subset a$. On voit donc que la condition O y est remplie, c'est-à-dire que l'espace est ordonné.

Terminons par une remarque concernant les espaces transitifs.

Théorème VIII. *Tout espace métrique presque ordonné qui est simplement ordonné, est nécessairement transitif.*

Démonstration. Supposons que, dans l'espace en question, un point c se trouve entre x et y et que x et y se trouvent entre a et b. Alors, en vertu du Théorème VII, on aura, par exemple,

$$x \subset c \subset y,$$
$$a \subset x \subset b,$$
$$a \subset y \subset b.$$

On en conclut, en vertu du Lemme VI, qu'on aura aussi

$$a \subset c \subset b.$$

Par suite, en vertu du Théorème VII, c se trouvera, lui-aussi, entre a et b. Ainsi, on voit que si le point c se trouve entre x et y et si x et y se trouvent entre a et b, le point c se trouve, lui-aussi, entre a et b. C'est la condition T, c'est-à-dire que l'espace est transitif.

APPENDICE.

Sur les diverses définitions des systèmes de choses.

Plusieurs travaux ont déjà été consacrés à l'étude des systèmes de choses. Outre un Mémoire de R. Dedekind,[3] on peut citer les recherches récentes de MM. K. Menger,[4] Fritz Klein,[5] Garrett Birkhoff[6] et O. Ore.[7] C'est M. Menger qui a proposé le nom System von Dingen. M. Klein l'appelle Verband, M. Birkhoff préfère le nom Lattice, M. Ore le nom Structure.

Au fond, il s'agit toujours de la même chose à quelques détails près. Il est à noter seulement que la forme de la définition du système de choses que j'ai adoptée dans le présent article diffère de celle qui a été adoptée par tous les auteurs cités sauf M. Ore.[8] Tandis que, pour moi et pour M. Ore, le point de départ est la relation $a \subset b$, les autres commencent par les opérations ab et $a + b$. Dans ce dernier cas, on construit le système d'axiomes pour le système de choses comme il suit:

1*. A tout couple de choses a, b correspond une chose déterminée ab telle que

$$(ab)c = a(bc), \quad ab = ba, \quad aa = a.$$

2*. A tout couple de choses a, b correspond une chose déterminée $a + b$ telle que

$$(a + b) + c = a + (b + c), \quad a + b = b + a, \quad a + a = a.$$

3*. $ab = a$ entraîne $a + b = b$, et inversement.

---

[3] "Über die von drei Moduln erzeugte Dualgruppe," *Mathematische Annalen*, vol. 53 (1900), pp. 371-403.

[4] "Axiomatik der endlichen Mengen und der elementargeometrischen Verknüpfungsbeziehungen," *Jahr. d. Deutsch. Math.-Ver.*, vol. 37 (1928), pp. 309-325.

[5] "Zur Theorie der abstrakten Verknüpfungen," *Mathematische Annalen*, vol. 105 (1931), pp. 308-323.

[6] "On the combination of subalgebras" et "Applications of lattice algebra," *Proceedings of the Cambridge Philosophical Society*, vol. 23 (1933), pp. 441-464, resp. vol. 30 (1934), pp. 115-122.

[7] "On the foundation of abstract algebra," I, *Annals of Mathematics*, vol. 36 (1935), pp. 406-437.

[8] Dans un article récent, "New foundations of projective and affine geometry," *Annals of Mathematics*, vol. 37 (1936), pp. 456-482, qui a paru après la rédaction de mon mémoire, K. Menger est parti du même point de vue que moi. Ses méthodes diffèrent et certains de ses résultats se dirigent dans une direction différente des miens.

Puis, on introduit la relation $a \subset b$ en la définissant comme équivalente à celle qui figure dans le dernier de ces axiomes, c'est-à-dire à $ab = a$ ou à $a + b = b$.

On y ajoute parfois l'axiome:

4*. Il existe une chose 0 telle qu'on a toujours

$$0 \subset z;$$

et parfois aussi l'axiome:

5*. Il existe une chose 1 telle qu'on a toujours

$$z \subset 1.$$

Il n'y a aucune difficulté à établir l'équivalence des axiomes 1*, 2*, 3*, et des axiomes 1°, 2°, 3° du $n^o 2$ du présent article.

Les systèmes satisfaisant aux axiomes 1*, 2*, 3* méritent le nom des systèmes de choses dans le sens le plus large. En y ajoutant l'axiome 4*, on obtient précisemment les systèmes de choses dont nous nous avons occupé. Enfin, en y ajoutant l'axiome 5*, on obtient les systèmes de choses dans le sens le plus étroit.

Je veux noter encore que certains raisonnements que nous avons employés, ont été utilisés déjà par les auteurs cités. Ainsi, on trouve chez M. Menger l'ensemble des sous-espaces linéaires d'un espace projectif comme un système de choses normées; et l'on trouve chez M. Birkhoff l'emploi de l'expression $(a + b)c + ab$ dans l'étude des propriétés caractéristiques des systèmes distributifs.

KLIAZMA, près de
MOSCOU, U. S. S. R.

A SYMMETRIC REDUCTION OF THE PLANAR THREE-BODY PROBLEM.

By F. D. MURNAGHAN.

Since the center of mass of a dynamical system, consisting of a collection of particles, in which the only forces are the mutual attractions of the various particles, remains at rest in an inertial reference frame, we shall suppose the center of mass of the three particles, of masses m_1, m_2, m_3, respectively, to be the origin of our inertial frame. The coördinates of the point P_k occupied by m_k being denoted by (x_k, y_k), $k = 1, 2, 3$, this implies the equations

$$(1) \qquad \Sigma\, m_k x_k = 0; \qquad \Sigma\, m_k y_k = 0$$

and the system is accordingly one with four degrees of freedom—there being six coördinates (x_k, y_k) connected by the two relations (1). The kinetic energy of the system is $T = \Sigma\, \frac{1}{2} m_k v_k^2$ where $v_k^2 = (\dot{x}_k^2 + \dot{y}_k^2)$ is the squared velocity of the k-th particle and the relations (1) enable us to readily express T in terms of the relative velocities of the three particles.[1] In fact, on squaring the relation $\Sigma\, m_k \dot{x}_k = 0$ and replacing each product term $\dot{x}_p \dot{x}_q$ by its equivalent $\frac{1}{2}[\dot{x}_p^2 + \dot{x}_q^2 - (\dot{x}_p - \dot{x}_q)^2]$, we find

$$(\Sigma\, m_k)(\Sigma\, m_k \dot{x}_k^2) = \Sigma\, m_p m_q (\dot{x}_p - \dot{x}_q)^2.$$

This implies

$$(2) \qquad 2T = (\Sigma\, m_p m_q v_{pq}^2)/\Sigma\, m_k$$

where $v_{pq}^2 = (\dot{x}_p - \dot{x}_q)^2 + (\dot{y}_p - \dot{y}_q)^2$ is the squared relative velocity of the particles m_p and m_q. We now denote by (a_1, a_2, a_3) and (A_1, A_2, A_3) the sides and angles of the triangle formed by the three particles and by $(\theta_1, \theta_2, \theta_3)$ the inclinations of the sides, whose lengths are (a_1, a_2, a_3) respectively, to the x-axis of our inertial frame (the sides having the sense found by traversing the triangle in the order $1 \to 2 \to 3 \to 1$). It is clear that $\theta_2 - \theta_3 \equiv A_1 \pmod{\pi}$ and hence $\dot{\theta}_2 - \dot{\theta}_3 = \dot{A}_1$; similarly $\dot{\theta}_3 - \dot{\theta}_1 = \dot{A}_2$, $\dot{\theta}_1 - \dot{\theta}_2 = \dot{A}_3$. These equations enable us to express each of the angular velocities $\dot{\theta}_k$ in terms of $\dot{\phi}$ and the $\dot{A}_k$ where $3\phi = \theta_1 + \theta_2 + \theta_3$. For instance

$$3\dot{\phi} = \dot{\theta}_1 + \dot{\theta}_2 + \dot{\theta}_3 = 3\dot{\theta}_3 + \dot{A}_1 - \dot{A}_2$$

[1] E. J. Routh, *Stability of Motion* (1877), p. 67.

so that $\dot{\theta}_3 = \dot{\phi} + (\dot{A}_2 - \dot{A}_1)/3$. Now the polar coördinates of P_2 relative to P_1 being (a_3, θ_3) we have

$$v_{21}^2 = \dot{a}_3^2 + a_3^2\dot{\theta}_3^2 = \dot{a}_3^2 + a_3^2\{\dot{\phi} + (\dot{A}_2 - \dot{A}_1)/3\}^2.$$

Similarly

$$v_{32}^2 = \dot{a}_1^2 + a_1^2\{\dot{\phi} + (\dot{A}_3 - \dot{A}_2)/3\}^2$$
$$v_{13}^2 = \dot{a}_2^2 + a_2^2\{\dot{\phi} + (\dot{A}_1 - \dot{A}_3)/3\}^2.$$

On substituting these values in the expression (2) for $2T$ we readily find

$$2T = R_2\dot{\phi}^2 + 2R_1\dot{\phi} + R_0$$

where the coefficients (R_2, R_1, R_0) are furnished by the formulae

$$\begin{aligned} \mu R_2 &= \Sigma\, m_2 m_3 a_1^2 \\ \mu R_1 &= \tfrac{1}{3}\Sigma\, m_1(m_3 a_2^2 - m_2 a_3^2)\dot{A}_1 \\ \mu R_2 &= \Sigma\, m_2 m_3 \dot{a}_1^2 + \tfrac{1}{9}\Sigma\, m_2 m_3 a_1^2(\dot{A}_3 - \dot{A}_2)^2 \end{aligned} \tag{3}$$

where $\mu = m_1 + m_2 + m_3$ and the summation sign implies the sum of the term written and the two others obtained from it by cyclic permutation of the labels 1, 2, 3. The coefficients (R_2, R_1, R_0) are easily expressible in terms of the sides (a_1, a_2, a_3) and their time derivatives. In fact the relation $2a_2a_3 \cos A_1 = a_2^2 + a_3^2 - a_1^2$ furnishes (on differentiation with respect to t and making use of the relations $a_2 = a_3 \cos A_1 + a_1 \cos A_3$, etc.) the relations $2\Delta\dot{A}_1 = a_1(\dot{a}_1 - \dot{a}_2 \cos A_3 - \dot{a}_3 \cos A_2)$ etc., where Δ is the area of the triangle $P_1P_2P_3$.

Regarding, now, the three sides (a_1, a_2, a_3) and the angle ϕ as the four coördinates of the dynamical system it is clear that ϕ is an ignorable coördinate with the momentum integral

$$c = \partial T/\partial\dot{\phi} = R_2\dot{\phi} + R_1 = (1/\mu)\Sigma\, m_2 m_3 a_1^2 \dot{\theta}_1. \tag{4}$$

It follows, by a reasoning similar to that already given for the kinetic energy T, that c is the angular momentum of the system about its center of mass. Thus if we multiply $\Sigma\, m_k x_k = 0$ by $\Sigma\, m_k \dot{y}_k = 0$ we find

$$\Sigma\, m_k^2 x_k \dot{y}_k + \Sigma\, m_p m_q (x_p \dot{y}_q + x_q \dot{y}_p) = 0$$

and on replacing $x_p\dot{y}_q + x_q\dot{y}_p$ by the equivalent expression

$$x_p\dot{y}_p + x_q\dot{y}_q - (x_p - x_q)(\dot{y}_p - \dot{y}_q)$$

we find

$$\mu\Sigma\, m_k x_k \dot{y}_k = \Sigma\, m_p m_q (x_p - x_q)(\dot{y}_p - \dot{y}_q).$$

On interchanging the variables x and y and subtracting we see that the angular momentum of the system about the center of mass is the quotient, by the total mass μ, of the expression $\Sigma\, m_p m_q k_{pq}$ where

$$k_{pq} = (x_p - x_q)(\dot{y}_p - \dot{y}_q) - (y_p - y_q)(\dot{x}_p - \dot{x}_q)$$

is the angular momentum (per unit mass) of the particle m_p relative to the particle m_q. In terms of the relative polar coördinates (a, θ), k_{23}, for example, is $a_1^2\dot{\theta}_1$ and hence the angular momentum of the system about its center of mass is the constant c of formula (4).

We now make use of the angular momentum integral to reduce the problem to one of three degrees of freedom in which the coördinates are the lengths (a_1, a_2, a_3) of the sides of the triangle formed by the three particles. The potential energy being

$$V = -\gamma\left(\frac{m_2 m_3}{a_1} + \frac{m_3 m_1}{a_2} + \frac{m_1 m_2}{a_3}\right)$$

(where γ is the gravitational constant) the Lagrangian function of the original problem is $L = T - V$ and the modified Lagrangian (in the sense of Routh) is $L^* = L - c\dot{\phi}$. The modified Hamiltonian function is

$$H^* = \left(\Sigma\, \dot{a}_k \frac{\partial L^*}{\partial \dot{a}_k}\right) - L^*$$

and since

$$\frac{\partial L^*}{\partial \dot{a}_k} = \frac{\partial L}{\partial \dot{a}_k} + \frac{\partial L}{\partial \dot{\phi}}\frac{\partial \dot{\phi}}{\partial \dot{a}_k} - c\frac{\partial \dot{\phi}}{\partial \dot{a}_k} = \frac{\partial L}{\partial \dot{a}_k} \quad \left(\text{since } \frac{\partial L}{\partial \dot{\phi}} = \frac{\partial T}{\partial \dot{\phi}} = c\right)$$

we have

$$H^* = \left(\Sigma\, \dot{a}_k \frac{\partial L}{\partial \dot{a}_k}\right) + c\dot{\phi} - L = \left(\Sigma\, \dot{a}_k \frac{\partial L}{\partial \dot{a}_k}\right) + \frac{\partial L}{\partial \dot{\phi}}\dot{\phi} - L = H.$$

Hence, in order to find the modified Hamiltonian function, we have merely to express the original energy function $H = T + V$ in terms of the modified momenta $(\omega_1, \omega_2, \omega_3) = (\partial L^*/\partial \dot{a}_1, \partial L^*/\partial \dot{a}_2, \partial L^*/\partial \dot{a}_3)$. This is readily done if we observe that $2T$ is the sum of the squared magnitudes of the linear momentum vectors of the three masses each divided by the corresponding mass. The x-component of the linear momentum of the particle m_1 is

$$m_1\dot{x}_1 = \frac{\partial L}{\partial \dot{x}_1} = \left(\Sigma \frac{\partial L}{\partial \dot{a}_k}\frac{\partial \dot{a}_k}{\partial \dot{x}_1}\right) + \frac{\partial L}{\partial \dot{\phi}}\frac{\partial \dot{\phi}}{\partial \dot{x}_1} = \left(\Sigma\, \omega_k \frac{\partial \dot{a}_k}{\partial \dot{x}_1}\right) + c\frac{\partial \dot{\phi}}{\partial \dot{x}_1}.$$

But $\partial \dot{a}_k/\partial \dot{x}_1 = \partial a_k/\partial x_1$ and so is zero if $k = 1$, whilst if $k = 2$ or 3 it is the

x-direction cosine of the corresponding side of the triangle (with the sense *towards* the vertex P_1). Furthermore the relation

$$a_2{}^2\dot{\theta}_2 = (x_3 - x_1)(\dot{y}_3 - \dot{y}_1) - (y_3 - y_1)(\dot{x}_3 - \dot{x}_1)$$

yields

$$a_2 \frac{\partial \dot{\theta}_2}{\partial \dot{x}_1} = (y_3 - y_1)/a_2; \qquad a_2 \frac{\partial \dot{\theta}_2}{\partial \dot{y}_1} = -(x_3 - x_1)/a_2$$

so that the vector whose components are $(\partial\dot{\theta}_2/\partial\dot{x}_1, \partial\dot{\theta}_2/\partial\dot{y}_1)$ has magnitude $1/a_2$ and is $90°$ ahead of the side P_3P_1. Similarly for $(\partial\dot{\theta}_3/\partial\dot{x}_1, \partial\dot{\theta}_3/\partial\dot{y}_1)$ whilst $(\partial\dot{\theta}_1/\partial\dot{x}_1, \partial\dot{\theta}_1/\partial\dot{y}_1)$ is the zero vector. Hence the vector $(c\,\partial\dot{\phi}/\partial\dot{x}_1, c\,\partial\dot{\phi}/\partial\dot{y}_1)$ is the sum of two vectors of magnitudes $c/3a_2$ and $c/3a_3$ and $90°$ ahead of the sides P_3P_1 and P_2P_1 respectively. Hence the linear momentum of the particle m_1 may be analysed into the sum of four vectors: (1) a vector, of magnitude ω_2, along P_3P_1; (2) a vector, of magnitude ω_3, along P_2P_1; (3) a vector, of magnitude $c/3a_2$, $90°$ ahead of P_3P_1; and (4) a vector, of magnitude $c/3a_3$, $90°$ ahead of P_2P_1. Hence the squared magnitude of the linear momentum of the particle m_1 is

$$\omega_2{}^2 + \omega_3{}^2 + 2\omega_2\omega_3 \cos A_1 + \frac{c^2}{9}\left(\frac{1}{a_2{}^2} + \frac{1}{a_3{}^2}\right) + \tfrac{2}{9}\frac{c^2}{a_2a_3}\cos A_1 + \tfrac{2}{3}c \sin A_1\left(\frac{\omega_2}{a_3} - \frac{\omega_3}{a_2}\right)$$

and it follows that the Hamiltonian function of the reduced problem is

$$H^* = H = \Sigma \frac{1}{2m_i}\left[\omega_j{}^2 + \omega_k{}^2 + 2\left(\omega_j\omega_k + \frac{c^2}{9a_ja_k}\right)\cos A_i + \frac{c^2}{9}\left(\frac{1}{a_j{}^2} + \frac{1}{a_k{}^2}\right) + \tfrac{2}{3}\, c \sin A_i\left(\frac{\omega_j}{a_k} - \frac{\omega_k}{a_j}\right)\right] + V$$

where (i, j, k) is a cyclic arrangement of the labels $(1, 2, 3)$ (i. e. $= (1, 2, 3)$ or $(2, 3, 1)$ or $(3, 1, 2)$).[2]

THE JOHNS HOPKINS UNIVERSITY.

---

[2] References to previous reductions of the plane three body problem may be found in Whittaker, *Enzykl. d. Math. Wiss.*, Bd. 6, 2[1] (Art. 12). In addition the reader is referred to the article by Grammel, *Handbuch der Physik*, Bd. 5, pp. 346-349. The present reduction is a result of a discussion with our colleague, Professor Wintner, who has applied it effectively to the "regularisation" of Levi-Civita.

ON A GENERALIZED TANGENT VECTOR. II.[1]

By H. V. Craig.

1. *Introduction.* In this paper we proceed with certain of the ideas and results given in a previous paper.[2] In particular, we define by means of the tangent vector T_r two metric tensors and develop from each of them a connection. These connections give rise to a scheme of parallel displacement enjoying the following properties: (a) the auto-parallel curves are the extremals associated with F; (b) the scalar product of any two vectors undergoing parallel displacement is constant.

2. *Notation.* In addition to the notation employed in I, we shall introduce the operator ${}^vO_r$ defined by[3]

$$(2.1) \qquad {}^vO_r F = \sum_{u=v}^{m} (-1)^u {}_uC_v F_{(u)r}^{\cdot\cdot(u-v)}$$

F being any sufficiently differentiable function of the coördinates and their derivatives with respect to a parameter t up to and including order m. The quantities T_r, defined by

$$(2.2) \qquad T_r = -{}^1O_r F,$$

were adopted in I as the components of the tangent vector. Furthermore, wherever it is necessary to stress the fact that the highest order of derivative occurring in F is m, we shall write $F(m)$, and for the associated tangent vector $T_r(m)$.

3. *Some properties of T_r and related vectors.* It is well known that the invariance of $\int F(1)\,dt$ under a parameter change implies the identity

$$(3.1) \qquad x'^r T_r(1) = F(1),$$

[1] Presented to the American Mathematical Society, October 27, 1934.

[2] *American Journal of Mathematics*, vol. 57 (1935), p. 457. This paper will be referred to as I.

[3] The quantities ${}^vO_r$ were first found by J. L. Synge who established their vector character. See J. L. Synge, "Some intrinsic and derived vectors in a Kawaguchi space," *American Journal of Mathematics*, vol. 57 (1935), p. 679. Soon after the writer encountered them he received from Professor Synge a copy of the abstract of his paper.

since $T_r(1) = \ddot{F}_{(1)r}(1)$. Also, it was shown in I that if $\int F(m)\,dt$ is invariant under a transformation of parameter, then the following is an identity:

$$x'^r\, T_r(m) = F(m), \qquad (m = 1, 2, \cdots). \tag{3.2}$$

In addition, by performing the indicated operations, one may verify that

$$x'^r\, {}^1O_r\, F^2(1) = -2\, F^2(1), \tag{3.3}$$

and that

$${}^1O_r\,[F(1)'] = -\,{}^0O_r\, F(1). \tag{3.4}$$

And thus we are led to investigate the more general expressions

$$x'^r\, {}^1O_r\, F^2(m) \quad \text{and} \quad {}^vO_r[F(m)'].$$

As a first step, we shall demonstrate the formula

$${}^1O_r\, F^2 = 2 \sum_{w=0}^{m-1} (w+1) F^{(w)} \cdot {}^{w+1}O_r\, F. \tag{3.5}$$

Now, by virtue of the rule for differentiating a product,

$$(F\, F_{(u)r})^{(u-1)} = \sum_{w=0}^{u-1} {}_{u-1}C_w\, F^{(w)}\, \dot{F}^{(u-1-w)}_{(u)r}; \tag{3.6}$$

also, from (2.1),

$${}^1O_r\, F^2 = \sum_{u=1}^{m} (-1)^u\, u[(F^2)_{(u)r}]^{(u-1)}. \tag{3.7}$$

Hence, we may write

$${}^1O_r\, F^2 = 2 \sum_{u=1}^{m} (-1)^u \cdot u \cdot \sum_{w=0}^{u-1} {}_{u-1}C_w\, F^{(w)}\, \ddot{F}^{(u-1-w)}_{(u)r}. \tag{3.8}$$

If, in this, we first change the order of summation and then replace $u \cdot {}_{u-1}C_w$ with $(w+1) \cdot {}_uC_{w+1}$, we obtain the equalities

$$\begin{aligned} {}^1O_r\, F^2 &= 2 \sum_{w=0}^{m-1} F^{(w)} \sum_{u=w+1}^{m} (-1)^u\, u \cdot {}_{u-1}C_w\, \ddot{F}^{(u-w-1)}_{(u)r} \\ &= 2 \sum_{w=0}^{m-1} (w+1) F^{(w)} \sum_{u=w+1}^{m} (-1)^u\, {}_uC_{w+1}\, \ddot{F}^{(u-w-1)}_{(u)r} \\ &= 2 \sum_{w=0}^{m-1} (w+1) F^{(w)}\, {}^{w+1}O_r\, F, \end{aligned} \tag{3.9}$$

and (3.5) is established.

Remark. Examining (3.5), we observe that, if $x'^r \cdot {}^cO_r\, F(m) = -\delta_1{}^c F(m)$, the first of the expressions to be investigated $x'^r\, {}^1O_r\, F^2(m)$ reduces to $-2\, F^2(m)$. As a matter of fact, if $\int F(m)\,dt$, $(m > 1)$ is invariant x'^r is

normal[4] to each of ${}^{m}O_r F$ and ${}^{0}O_r F$ while $x'^{r}\,{}^{1}O_r F$, as we know, reduces to $-F$. Thus we are led to consider the following theorem:

THEOREM 3.1. *If* $\int F(m)\,dt$ *is invariant under an admissible parameter transformation, then* $x'^{r}\,{}^{c}O_r F = -\delta_1{}^{c} F$, $(c = 0, 1, \cdots, m)$.

Proof. Since

$$L_a = \sum_{u=a}^{m} {}_uC_a\, x^{(u-a+1)r} F_{(u)r} = \delta_a{}^1 F,^{5,6} \tag{3.10}$$

it follows at once that for $c > 1$, L_{a+c} and $\sum_{a=0}^{m-c} (-1)^{a+c}\,{}_{a+c}C_c.\ \ L_{a+c}{}^{(a)}$ vanish.

Therefore, since the special cases of the theorem $c = 0$, $c = 1$ have been demonstrated, it will suffice to establish that for c larger than one

$$\sum_{a=0}^{m-c} (-1)^{a+c}\,{}_{a+c}C_c\, L_{a+c}{}^{(a)} = x'^{r}\,{}^{c}O_r F \tag{3.11}$$

is an identity.

Dropping the index r from x and F, and employing first the definition of L_a and then the rule for differentiating a product, we obtain the set of equalities

$$\begin{aligned} L_{a+c}{}^{(a)} &= \sum_{u=a+c}^{m} {}_uC_{a+c}\left[x^{(u-a-c+1)} F_{(u)}\right]^{(a)} \\ &= \sum_{u=a+c}^{m} {}_uC_{a+c} \sum_{v=0}^{a} {}_aC_v x^{(u-a-c+v+1)} F_{(u)}^{\cdot\,\cdot\,(a-v)} \end{aligned} \tag{3.12}$$

and, replacing the umbral index u with $u + a + c$, the relationship

$$L_{a+c}{}^{(a)} = \sum_{u=0}^{m-(a+c)} {}_{u+a+c}C_{a+c} \sum_{v=0}^{a} {}_aC_v x^{(u+v+1)} F_{(u+a+c)}^{\cdot\,\cdot\,(a-v)}. \tag{3.13}$$

Substituting (3.13) in (3.11), we find that the coefficient of the free x' in the left member of (3.11) is

$$\sum_{a=0}^{m-c} (-1)^{a+c}\,{}_{a+c}C_c F_{(a+c)}{}^{(a)},$$

---

[4] See H. V. Craig, "On the solution of the Euler equations for their highest derivatives," *Bulletin of the American Mathematical Society*, vol. 14 (1930), p. 559.

[5] See I, p. 461.

[6] For a more general discussion of the invariance of integrals under a parameter transformation, reference may be made to Théophile de Donder, *Théorie invariantive du calcul des variations*, Paris (1935), pp. 53-57. De Donder's exposition is based on the work of R. Deladrière and J. Géhéniau and treats of multiple integrals of functions containing higher order derivatives. My attention was called to it by a referee.

which, upon replacing a with $a - c$, becomes

$$\sum_{a=c}^{m} (-1)^a \, {}_aC_c F_{(a)}{}^{(a-c)},$$

the coefficient of x' in the right member. Likewise, if we put $v = 0$, $u = 1$ and then $u = 0$, $v = 1$, we obtain the coefficient of x'',

$$\sum_{a=0}^{m-c-1} (-1)^{a+c} \, {}_{a+c}C_c \cdot {}_{1+a+c}C_{a+c} \, {}_aC_0 F_{(a+c+1)}{}^{(a)} + \sum_{a=1}^{m-c} (-1)^{a+c} \, {}_{a+c}C_c \, {}_{a+c}C_{a+c} \, {}_aC_1 F_{(a+c)}{}^{(a-1)}$$

or

$$\sum_{w=0}^{1} \sum_{a=w}^{m-c-1+w} (-1)^{a+c} \, {}_{a+c}C_c \cdot {}_{a+c-w+1}C_{a+c} \cdot {}_aC_w F_{(a+c+1-w)}{}^{(a-w)}.$$

In general, the coefficient of $x^{(e+1)}$ is

$$\sum_{w=0}^{e} \sum_{a=w}^{m-c-e+w} (-1)^{a+c} \, {}_{a+c}C_c \cdot {}_{a+c+e-w}C_{a+c} \cdot {}_aC_w F_{(a+c+e-w)}{}^{(a-w)}$$

or, replacing a with $a + w$,

$$\sum_{w=0}^{e} \sum_{a=0}^{m-c-e} (-1)^{a+w+c} \, {}_{a+w+c}C_c \cdot {}_{a+c+e}C_{a+w+c} \cdot {}_{a+w}C_w F_{(a+c+e)}{}^{(a)}.$$

But, for $e > 0$ the foregoing expression vanishes, since

$$\sum_{w=0}^{e} (-1)^{a+w+c} \frac{(a+w+c)!\,(a+c+e)!\,(a+w)!}{(a+w)!\,c!\,(e-w)!\,(a+w+c)!\,a!\,w!} = \frac{(a+c+e)!}{a!\,c!\,e!} (-1)^{a+c} \sum_{w=0}^{e} (-1)^w \, {}_eC_w,$$

and the theorem is established. We now turn to the generalization of (3.3).

THEOREM 3.2. *If $\int F(m)\,dt$ is invariant under an admissible parameter transformation, then*

$$x'^r \, {}^1O_r \, F^2(m) = -2F^2(m).$$

This theorem is an immediate consequence of (2.2), (3.2), (3.5), and Theorem 3.1.

A generalization of Theorem 3.2 is as follows:

THEOREM 3.3. *If $\int F(m)\,dt$ is invariant under an admissible parameter transformation, then*

$$x'^r \cdot {}^cO_r \, \phi(F) = -[\delta_0{}^c (d\phi/dF)' + \delta_1{}^c d\phi/dF]F \qquad (c = 0, 1, \cdots, m).$$

Proof. By employing the definition of ${}^{c}O_r$ and the rule for differentiating a product and then changing the order of summation, we derive the set of equalities

$$
\begin{aligned}
(3.14)\qquad x'^r\,{}^{c}O_r\,\phi(F) &= x'^r \sum_{u=c}^{m} (-1)^u\,{}_uC_c\,[d\phi/dF\,F_{(u)r}]^{(u-c)} \\
&= x'^r \sum_{u=c}^{m} (-1)^u\,{}_uC_c \sum_{w=0}^{u-c} {}_{u-c}C_w\,(d\phi/dF)^{(w)} F_{(u)r}^{(u-c-w)} \\
&= x'^r \sum_{w=0}^{m-c} (d\phi/dF)^{(w)} \sum_{u=w+c}^{m} (-1)^u\,{}_uC_c\,{}_{u-c}C_w\,F_{(u)r}^{(u-c-w)}.
\end{aligned}
$$

But ${}_uC_c \cdot {}_{u-c}C_w = {}_{c+w}C_c \cdot {}_uC_{w+c}$, hence the last member of (3.14) may be expressed in the form

$$x'^r \sum_{w=0}^{m-c} {}_{c+w}C_c (d\phi/dF)^{(w)} \cdot {}^{w+c}O_r\,F,$$

and our proposition follows by way of Theorem 3.1. We now turn to the generalization of the identity (3.4).

THEOREM 3.4. *If $F(m)$ is any function such that ${}^{v}O_r\,F'$ exists, and if ${}^{-1}O_r\,F$ represents zero, then ${}^{v}O_r\,F' = -\,{}^{v-1}O_r\,F$, $(v = 0, 1, \cdots, m)$.*

Proof. If we define the symbol $F_{(-1)r}$ to be zero then we may write

$$(3.15)\qquad F' = \sum_{u=0}^{m} x^{(u+1)r} F_{(u)r}; \qquad F'_{(u)r} = F_{(u)r}{}' + F_{(u-1)r}$$

and, dropping the index r,

$$(3.16)\quad {}^{v}O\,F' = \sum_{u=v}^{m+1} (-1)^u\,{}_uC_v\,F'^{(u-v)}_{(u)} = \sum_{u=v}^{m+1} (-1)^u\,{}_uC_v[F_{(u)}{}' + F_{(u-1)}]^{(u-v)}.$$

Evidently, we may replace $m+1$ with m in the first term of the last member of (3.16), and the index u with $u+1$ in the second. The result is

$$(3.17)\quad {}^{v}O\,F' = -\sum_{u=v}^{m} (-1)^u[-\,{}_uC_v + {}_{u+1}C_v]F_{(u)}{}^{(u-v+1)} - (-1)^{v-1}F_{(v-1)};$$

and this, since ${}_{u+1}C_v - {}_uC_v = {}_uC_{v-1}$, reduces to $-\,{}^{v-1}O\,F$.

Another property of the Finsler tangent vector $-\,{}^{1}O_r\,F(1)$ or $F_{(1)r}$ due to Cartan,[7] may be stated as follows: if $x = x(s, t)$ is a parametric representation of a tube, such that, for each fixed s, $x = x(s, t)$ is an extremal while

[7] See E. Cartan, *Leçons sur les invariants intégraux*, Paris, Hermann (1922), chapters 1 and 18.

12

for each fixed t, $x = x(s, t)$ is a closed curve and if (varying our notation) $F_{(1)r}(e)$ is the covariant tangent vector to the extremal through the point in question, then the integral $\int F_{(1)r}(e)\,(\partial x^r/\partial s)\,ds$, taken around an s-curve, i. e. a curve on which t is constant, is independent of t. Unfortunately, this property does not carry over in full strength to $F(m)$.

As a first step in the investigation of this matter we shall demonstrate that, if V^r is any vector and $V^{(u)r}$ denotes d^uV^r/dt^u, then

$$(3.18)\quad \sum_{w=1}^{m}(-1)^w(V^r\cdot {}^wO_r F)^{(w-1)} = \sum_{v=1}^{m}(-1)^{v-1}\sum_{u=0}^{m-v} V^{(u)r}F_{\substack{(u+v)r}}^{\cdot\cdot(v-1)}$$

is an identity. This may be accomplished by comparing the corresponding coefficients of the derivatives of V^r. Thus, suppressing the index r, the coefficient of $V^{(a)}$ in the right member of (3.18), which we shall represent by CR, is given by

$$(3.19)\qquad CR = \sum_{v=1}^{m-a}(-1)^{v-1}F_{(a+v)}^{\cdot\cdot(v-1)}$$

Whereas, since the left member of (3.18) may be written in the form

$$\sum_{w=1}^{m}(-1)^w\sum_{u=0}^{w-1}{}_{w-1}C_u\,V^{(u)}({}^wOF)^{(w-1-u)},$$

the corresponding coefficient, CL, is

$$\sum_{w=a+1}^{m}(-1)^w\,{}_{w-1}C_a({}^wOF)^{(w-1-a)}.$$

Upon expanding ${}^wOF$ in CL and changing the order of summation, we obtain successively

$$(3.20)\qquad CL = \sum_{w=a+1}^{m}(-1)^w\,{}_{w-1}C_a\sum_{v=w}^{m}(-1)^v\,{}_vC_w\,F_{(v)}^{\cdot(v-1-a)}$$

$$= \sum_{v=a+1}^{m}\sum_{w=a+1}^{v}(-1)^{w+v}\,{}_{w-1}C_a\cdot{}_vC_w\,F_{(v)}^{\cdot(v-1-a)}$$

and, replacing v with $v+a$ and w with $w+a+1$,

$$(3.21)\qquad CL = \sum_{v=1}^{m-a}\sum_{w=0}^{v-1}(-1)^{w+v+1}\,{}_{w+a}C_a\;{}_{v+a}C_{w+a+1}\,F_{(a+v)}^{\cdot\cdot(v-1)}.$$

Evidently, $CL = CR$ if

$$(3.22) \qquad \sum_{w=0}^{v-1} (-1)^w \, {}_{w+a}C_a \cdot {}_{v+a}C_{w+a+1} = 1$$

or, expressed otherwise, if

$$(3.23) \qquad \sum_{w=0}^{v-1} (-1)^w \, {}_{v-1}C_w/(w+a+1) = a!\,(v-1)!/(v+a)!$$

If in the left member of (3.22) we set $a=0$, replace w with $w-1$ and subtract unity there results the expression $-\sum_{w=0}^{v} (-1)^w \, {}_vC_w$ which, as we know, vanishes. Therefore, (3.23) is valid for each integer v if $a=0$, and we proceed by induction. Specifically, we show that if for some a, (3.23) is an identity in V, then (3.23) is an identity in V for $a+1$. Thus, substituting $a+1$ for a in the left member of (3.23), we obtain

$$\sum_{w=0}^{v-1} (-1)^w \, {}_{v-1}C_w/(w+a+2).$$

But, upon replacing w with $w-1$ and performing certain simple operations, this expression may be transformed as follows:

$$\begin{aligned}(3.24) \qquad \sum_{w=0}^{v-1} (-1)^w \, {}_{v-1}C_w/(w+a+2) &= \sum_{w=1}^{v} (-1)^{w-1} \, {}_{v-1}C_{w-1}/(w+a+1) \\ &= -1/v \sum_{w=1}^{v} (-1)^w \, {}_vC_w \, w/(w+a+1) \\ &= -1/v \sum_{w=0}^{v} (-1)^w \, {}_vC_w \left(1 - \frac{a+1}{w+a+1}\right) \\ &= \frac{(a+1)}{v} \, \frac{a!\,v!}{(v+a+1)!},\end{aligned}$$

and the theorem (3.23) is established.

We shall now apply (3.18) to transform the first variation of $\int F dt$. Thus, let $x = x(s,t)$ be a surface such that for each fixed s, $x = x(s,t)$ is an extremal and let $J(s)$ represent $\int_{t_1}^{t_2} F dt$; then the first variation is given by (3.25)

$$(3.25) \qquad dJ/ds = \int_{t_1}^{t_2} \sum_{v=0}^{m} (\partial x^{(v)r}/\partial s) F_{(v)r} dt, \quad \text{where} \quad x^{(v)r} = \partial^v x^r/\partial t^v.$$

Now, by the usual integration by parts, (3.25) can be expressed, since ${}^0O_r\,F(e)$ vanishes, in the form

$$(3.26)\qquad dJ/ds = \sum_{v=1}^{m} [(-1)^{v-1} \sum_{u=v}^{m} F_{(u)r}(e)^{(v-1)} \partial x^{(u-v)r}/\partial s]_{t_2}^{t_1}.[8]$$

And finally, by virtue of (3.18) with $V^r = \partial x^r/\partial s$ and $u - v$ substituted for u, (3.25) takes the form

$$(3.27)\qquad dJ/ds = [T_r \partial x^r/\partial s]_{t_1}^{t_2} + [\sum_{w=2}^{m} (-1)^w (\partial x^r/\partial s \cdot {}^w O_r F(e))^{(w-1)}]_{t_1}^{t_2}.$$

This accomplished, we are ready to consider the generalization of Cartan's theorem.

THEOREM 3.5. *If $x = x(s,t)$ is an extremal tube then*

$$\sum_{w=1}^{m} (-1)^w [\int \{{}^w O_r F(e)\} (\partial x^r/\partial s)\, ds]^{(w-1)}$$

taken around an s-curve is independent of t.

Proof. The integral of dJ/ds around an s-curve vanishes and the theorem follows.

COROLLARY. *If $\sum_{w=2}^{m} (-1)^w (\partial x^r/\partial s) \cdot {}^w O_r F)^{(w-1)}$ vanishes over an extremal tube then $\int T_r (\partial x^r/\partial s)\, ds$ taken around an s-curve is independent of t.*

Now let us consider an extremal surface $x = x(s,t)$, not necessarily a tube, and let u represent s on an s-curve and t on an extremal or t-curve. If, in addition, $F dt$ is invariant under the transformation $t = t(T)$, the integral around a parameter mesh: s_1, t_1; s_2, t_2 of the quantity $S du$,

$$(3.28)\qquad S = \sum_{w=1}^{m} [(\partial x^r/\partial u)(-1)^w\, {}^w O_r F(e)]^{(w-1)},$$

vanishes. For, if $J(s)$ denotes the integral $\int_{t_1}^{t_2} F(s,t)\,dt$, then, since S reduces to F when u becomes t, the difference $\int_{t_1}^{t_2} S(s_1,t)\,dt - \int_{t_1}^{t_2} S(s_2,t)\,dt$ is $J(s_1) - J(s_2)$, which may be written $-\int_{s_1}^{s_2} (dJ/ds)\,ds$. But dJ/ds is $S(s,t_2) - S(s,t_1)$ with u replaced by s.

This result may be extended to other closed curves: $s = s(u)$, $t = t(u)$ if we replace the notation for the tangent vector $\partial x^r/\partial u$ with dx^r/du,

---

[8] See J. L. Synge, *loc. cit.*, p. 682, equation 2.7.

$dx^r/du = (\partial x^r/\partial s)\,ds/du + x'^r dt/du$ and interpret $(dx^r/du)^{(w)}$ to be $[\partial x^{(w)r}/\partial s]\,ds/du + x^{(w+1)r}dt/du$.

To summarize, we have seen that the vector $T_r(m)$ is in some respects analogous to the covariant tangent vector $T_r(1)$. We shall now turn to the development of a scheme of parallel displacement.

4. *Parallel displacement.* We assume throughout this section that Fdt is invariant under the parameter transformation: $t = t(T)$, $T = T(t)$, and observe that this change of variable induces the following transformation:

$$
(4.1)\quad \begin{cases} x(t(T)) = X(T); \qquad dx/dt = (dX/dT)\,dT/dt; \\ d^m x/dt^m = (d^m X/dT^m)(dT/dt)^m + \cdots; \\ F(x, \cdots, x^{(m)}) = F(X, \cdots, d^m X/dT^m)\,dT/dt \text{ or } F(x) = F(X)\,dT/dt; \\ dT/dt\, F_{(m)r}(X) = F_{(m)r}(x)\cdot(dT/dt)^m; \end{cases}
$$

and, therefore, $F^{2m-1}F_{(m)r(m)s}$ is invariant. Furthermore, if $t = aT + b$, then

$$
(4.2)\qquad F_{(w)r}(X) = F_{(w)r}(x)\cdot(1/a)^{w-1}
$$

and consequently $F^{\cdot\cdot(w-1)}_{(w)r}$ transforms by invariance and likewise T_r. Hence, we shall suppose in what follows that the parameter is linearly related to the arc length and adopt the quantities f_{rs},

$$
(4.3)\qquad f_{rs} = F^{2m-1}F_{(m)r(m)s} + T_r T_s,
$$

as the components of our metric tensor.

Following the procedure outlined in I, we construct an auxiliary connection and then modify it by means of the vectors T_r and S_r. As a first step we differentiate the equation of transformation of the fundamental tensor, thus

$$
(4.4)\qquad \bar{f}_{ij}' = f_{rs}'X_i^r X_j^s + f_{rs}X_i^{r\prime}X_j^s + f_{rs}X_i^r X_j^{s\prime}.
$$

Then recalling the relationships [9]

$$
(4.5)\qquad X^{(m-1)s}_{(m-1)j} = X_j^s; \qquad X^{(m)s}_{(m-1)j} = mX'^s_j = mX_j^{s\prime},
$$

and the tensor character of $F_{(m)r}$ and T_r, we derive the transformations:

$$
(4.6)\qquad \bar{F}_{(m)i(m-1)j} = F_{(m)r(m-1)s}X_i^r X_j^s + F_{(m)r(m)s}X_i^r m X_j^{s\prime};
$$

---

[9] See I, p. 457.

$$(4.7) \qquad \bar{T}_j' = T_s' X_j^s + T_s X_j^{s\prime};$$

$$(4.8) \qquad \begin{aligned} &\bar{F}^{2m-1} \bar{F}_{(m)i(m-1)j} + m\bar{T}_i \bar{T}_j' \\ &\quad = (F^{2m-1} F_{(m)r(m-1)s} + mT_r T_s') X_i^r X_j^s + m f_{rs} X_i^r X_j^{s\prime}. \end{aligned}$$

If we permute the indices i and j in (4. 8), subtract the result from (4. 8) and indicate this operation of permutation and subtraction in the well-known way by means of brackets, we may write

$$(4.9) \qquad \begin{aligned} &\bar{F}^{2m-1} \bar{F}_{(m)[i(m-1)j]} + m\bar{T}_{[i} \bar{T}_{j]}' \\ &\quad = (F^{2m-1} F_{(m)[r(m-1)s]} + mT_{[r} T_{s]}') X_i^r X_j^s + m f_{rs} X^r{}_{[i} X^s{}_{j]}'. \end{aligned}$$

Multiplying this last equation by $1/m$ and subtracting the result from (4. 4), we obtain our auxiliary connection.

$$(4.10) \qquad \begin{aligned} \overline{\{i, j\}} &= \{r, s\} X_i^r X_j^s + f_{rs} X_i^{r\prime} X_j^s, \\ 2\{r, s\} &= f_{rs}' - [1/m F^{2m-1} F_{(m)[r(m-1)s]} + T_{[r} T_{s]}']. \end{aligned}$$

Assuming that F is positive, the determinant $| F_{(m)r(m)s} |$ of rank $n-1$ and that the "F one" function of the calculus of variations is non-vanishing, the determinant [10] $| f_{rs} |$ is likewise non-vanishing and the normalized cofactors f^{rs} of $| f_{rs} |$ exist. Consequently, we may raise the second index j in (4. 10) by multiplying by $Y_t{}^j f^{tu} = \bar{f}^{jm} X_m{}^u$ thus,

$$(4.11) \qquad X_m{}^u \overline{\left\{ {m \atop i} \right\}} = \left\{ {u \atop r} \right\} X_i^r + X_i^{u\prime}; \qquad \left\{ {u \atop r} \right\} = f^{tu} \{r, t\}.$$

This accomplished, the formula for $\theta T^{r\cdots}_{s\cdots}$ and the theorems concerning the derivatives of sums, products, the Kronecker delta, and scalars expressed by contraction of tensors may be obtained as in Finsler geometry.

Furthermore, from the definitions of $\{r, s\}$ and $\left\{ {u \atop r} \right\}$, we observe at once that $\{r, s\} + \{s, r\} = f_{rs}'$, $f_{us} \left\{ {u \atop r} \right\} = \{r, s\}$ and consequently, by expanding θf_{rs}, we have the relationship

$$(4.12) \qquad \theta f_{rs} = f_{rs}' - f_{us} \left\{ {u \atop r} \right\} - f_{ru} \left\{ {u \atop s} \right\} = 0.$$

Obviously, if we alter $\{r, s\}$ by adding a skew symmetric tensor, the above equality will hold as before. With this in mind, we introduce our second connection $\{r, s\}^*$, $\left\{ {t \atop r} \right\}^*$ defined as follows:

---

[10] See H. V. Craig, "On the solution of the Euler equations for their highest derivatives," *Bulletin of the American Mathematical Society*, vol. 36 (1930), p. 562.

$$
(4.13)\qquad
\begin{aligned}
&\{r, s\}^* = \{r, s\} + S_r T_s - T_r S_s,\\
&\{{}^t_r\}^* = f^{st}\{r, s\}^* = \{{}^t_r\} + f^{st} T_s S_r - T_r S^t,\\
&\Big(S_r = F_{(0)r} + \sum_{u=2}^{m} (u-1)(-1)^{u-1} F_{(u)r^{(u)}} - T_t\{{}^t_r\}\Big).
\end{aligned}
$$

Let us next examine the vector $\theta^* T_r$, which is to be computed, as the presence of the asterisk suggests, by means of the new connection. Thus,

$$
(4.14)\qquad \theta^* T_r = T_r' - T_t\{r\} - f^{st} T_s T_t S_r + T_r T_t S^t.
$$

By virtue of the invariance of $F dt$ we have the equalities: $x'^r T_r = F$, $x'^r F_{(m)r(m)s} = 0$ and, therefore, $x'^r f_{rs} = F T_s$. Consequently, we may write $x'^t = x'^r f_{rs} f^{st} = F T_s f^{st}$ and, multiplying by T_t, $F = F f^{st} T_s T_t$. That is, anticipating a definition, T_s is a unit vector regardless of the parameter. Hence, since E_r is $\theta T_r - S_r$, $\theta^* T_r$ reduces to $E_r + T_r S_t x'^t / F$. But, again by virtue of the invariance of $F dt$, $x'^t E_t$ vanishes [11] and $x'^t S_t$ becomes $x'^t \theta T_t$. Furthermore, due to the vanishing of θf_{rs} and $\theta \delta_s{}^r$, we conclude that θf^{rs} is likewise zero and the following holds:

$$
(4.15)\qquad 2(\theta T_r) x'^r / F = 2(\theta T_r) f^{rs} T_s = \theta(f^{rs} T_r T_s) = 0;
$$

hence

$$
(4.16)\qquad \theta^* T_r = E_r.
$$

Also, if the parameter is the generalized arc, then $x'^r f_{rs}$ is T_s and, therefore, if we write E^t for $f^{st} E_s$, it follows that $\theta^* x'^t$ is E^t.

Finally, if we define the generalized magnitude of a vector, angle, and parallel displacement in the usual way, which may be indicated by the following identities and conditional equality

$$
(4.17)\qquad |V|^2 = f_{rs} V^r V^s, \quad |U|\,|V| \cos(U, V) = f_{rs} U^r V^s, \quad \theta^* V^r = 0,
$$

we have a theory of displacement possessing the cardinal properties of the Synge-Taylor parallelism.[12] Briefly these characteristics are: (a) the magnitudes of and angles between vectors undergoing parallel displacement are constant, and (b) the autoparallel curves are the extremals associated with F.

---

[11] See H. V. Craig, "On the solution of the Euler equations for their highest derivatives," *loc. cit.*, p. 560.

[12] Developed independently by J. L. Synge and J. H. Taylor. See J. L. Synge, "A generalization of the Riemannian line element," *Transactions of the American Mathematical Society*, vol. 27 (1925), pp. 61-67, and J. H. Taylor, "A generalization of Levi-Civita's parallelism and the Frenet formulas," *ibid.*, pp. 246-264.

Furthermore, relative to the foregoing connections, we may define derivatives [13] (covariant in character) of tensors whose components are functions of $x, x', \cdots, x^{(m)}$ by means of the identities

$$(4.18) \quad S_{,r} = S_{(m-1)r} - mS_{(m)s}\{{}^{s}_{r}\}; \quad T^r{}_{s,t} = T^r{}_{s(m-1)t} - mT^r{}_{s(m)u}\{{}^{u}_{t}\}.$$

A second metric tensor (from which we will develop a theory of parallelism) may be obtained quite naturally from considerations based on the interpretation of metric space given in I. Briefly, this view is as follows: We suppose that there is given a Euclidean $n+1$ space (x^r, z) together with the set of all arcs, $x^r = x^r(t)$ of class C^m lying in the subspace (x^r) and associate with each of these base arcs the warped arc,

$$x^r = x^r(t), \qquad z = \int_0^t \sqrt{F^2 - E_{rs}x'^r x'^s}\, dt + k.$$

The quantities E_{rs} are merely the components of the Euclidean metric tensor and k is a constant. The coördinate systems to be admitted are either rectangular Cartesian or those obtainable from such by transformation of the x's only. Hence, the element of arc is given by the equality

$$(ds/dt)^2 = E_{rs}x'^r x'^s + z'^2 = F^2.$$

By way of illustration let us suppose that we have given the surface $z = z(x^1, x^2)$. Then

$$\begin{aligned}(ds/dt)^2 &= E_{rs}x'^r x'^s + Z_1{}^2(x'^1)^2 + 2Z_1Z_2x'^1x'^2 + Z_2{}^2(x'^2)^2 \\ &= (E_{rs} + Z_rZ_s)x'^r x'^s = F^2, \qquad (Z_1 = \partial z/\partial x^1,\ Z_2 = \partial z/\partial x^2).\end{aligned}$$

It is of course well known that we can study the warped or surface arcs through their corresponding base curves $x^1(t)$, $x^2(t)$ if we replace the Euclidean metric E_{rs} with the Riemannian metric $E_{rs} + Z_rZ_s$. The functions F dealt with in this paper are not necessarily Riemannian and, hence, the warped curves need not lie on a surface.

Returning to the general case, we shall select as parameter the arc length $\bar{s}$ of the warped curve in question, consequently x'^r, z' is the unit tangent vector and F maintains the value unity along this curve. Evidently, if the arc is such that the quantities

[13] See H. V. Craig, "On a covariant differentiation process," *Bulletin of the American Mathematical Society*, vol. 37 (1931), pp. 731-734; also Paper II, *ibid.*, vol. 39 (1933), pp. 919-922.

(4.19) $$Z_r = (FT_r - E_{ra}x'^a)(F^2 - E_{ab}x'^a x'^b)^{-\frac{1}{2}}$$

exist, we may write

(4.20) $$x'^r x'^s(E_{rs} + Z_r Z_s) = E_{rs}x'^r x'^s + z'^2 = 1,$$

and are thus led to define f_{rs}, the metric tensor, to be $E_{rs} + Z_r Z_s$. We note in passing, distinguishing the warped curves from the base curves by means of a wave, that

(4.21) $$(F^2 - E_{ab}x'^a x'^b)^{\frac{1}{2}} = [1 - E_{ab}dx^a/ds\, dx^b/d\bar{s}\,(ds/d\bar{s})^2]^{\frac{1}{2}} = \sin(c, \bar{c}).$$

Furthermore, $x'^s(E_{rs} + Z_r Z_s) = FT_r = T_r$; that is, our generalized tangent vector is the covariant description of the vector x'^r. Similarly, we may raise indices by means of the normalized cofactors f^{rs} whose existence depends upon the non-vanishing of the determinant $|f_{rs}|$, a fact which may be established readily. For, by employing a rectangular Cartesian coördinate system and certain well known rules for evaluating determinants, we have at once

(4.22) $$|E_{rs} + Z_r Z_s| = 1 + \Sigma(Z_r)^2 \neq 0.$$

Now, if C_1 and C_2 are any two intersecting base arcs, we can select the k's so that the associated warped arcs will intersect. And if, further, these are restricted as above the angles of intersection will be given by

(4.23) $$\cos\theta = [E_{rs} + Z_r(1)Z_s(2)]x'^r(1)x'^s(2),$$

the symbols (1), (2) indicating the arcs from which the quantities involved are computed. Thus, in computing angles, the value of the metric tensor is determined by two curves.

To form a connection, we start, as before, by differentiating the equation of transformation of the fundamental tensor, thus:

(4.24) $$\bar{f}_{ij}' = f_{rs}'X_i^r X_j^s + f_{rs}X_i^{r\prime}X_j^s + f_{rs}X_i^r X_j^{s\prime}.$$

Designating $x'^t E_{rt}$ by E_r and differentiating the relationship $\bar{E}_i = E_r X_i^r$ with respect to y^j, we obtain the equation,

(4.25) $$\bar{E}_{i(0)j} = E_{r(0)s}X_i^r X_j^s + E_{rs}X_i^r X_j^{s\prime} + E_r X^r{}_{ij}.$$

To this we add the equality obtained by multiplying

(4.26) $$\bar{Z}_j' = Z_s' X_j^s + Z_s X_j^{s\prime} \quad \text{and} \quad \bar{Z}_i = Z_r X_i^r.$$

The result is

$$(4.27)\quad \bar{E}_{i(0)j} + \bar{Z}_i\bar{Z}_j' = (E_{r(0)s} + Z_rZ_s')X_i^rX_j^s + f_{rs}X_i^rX_j^{s\prime} + E_rX^r{}_{ij},$$

from which we derive

$$(4.28)\quad \bar{E}_{[i(0)j]} + \bar{Z}_{[i}\bar{Z}_{j]}' = (E_{[r(0)s]} + Z_{[r}Z_{s]}')X_i^rX_j^s + f_{rs}X^r{}_{[i}X^s{}_{j]}'.$$

Finally, by subtracting (4.28) from (4.24) and dividing by two, we obtain a connection

$$(4.29)\quad \overline{\{i, j\}} = \{r, s\}X_i^rX_j^s + f_{rs}X_i^{r\prime}X_j^s,\quad 2\{r, s\} = f_{rs}' - E_{[r(0)s]} - Z_{[r}Z_{s]}',$$

which is such that, in the case of the Riemannian geometry discussed, $\{r, s\}$ is $\{rt, s\}x'^t$.

By means of the procedure applied to equation (4.10), we could develop a second theory of parallelism having the properties asserted of the first.

Texas University.

THE FUNDAMENTAL THEOREM FOR RIEMANN INTEGRALS.

By E. R. van Kampen.

1. The fundamental theorem for Riemann integrals states that a function defined on a closed bounded interval has an integral if and only if the function is bounded and its discontinuity points form a 0-set. The proofs of this theorem, as given in the literature,[1] still follow almost exactly the historical development. The upper and lower Darboux sums are defined and the condition that the upper and lower integrals of the function, defined by means of these sums, are equal is transformed into the condition that the discontinuity points of the function form a 0-set. In this note a shorter treatment will be given. This treatment is closely related to the usual treatment in case of a continuous function and is independent of the Darboux sums. The desirability of an investigation of the Riemann integral along these lines was pointed out to me by Wintner. The same method may be applied to n-dimensional Riemann integrals and to Riemann-Stieltjes integrals.

2. A *partition* Δ of the interval $a \leqq x \leqq b$ of the real variable x is a finite number of values $x_0 = a, x_1, \cdots, x_n = b$ such that $x_{i-1} < x_i$, $(i = 1, \cdots, n)$. The segments of the partition Δ are the intervals $x_{i-1} \leqq x \leqq x_i$ of the variable x. The *degree of fineness of* Δ is the maximum of the lengths of the segments of Δ and will be denoted by $d(\Delta)$. An intermediate partition Γ of Δ is a set of n values y_i of the variable x such that $x_{i-1} \leqq y_i \leqq x_i$.

3. Let $\{I_n\}$ be a sequence of open intervals I_n of the variable x. The length of I_n being denoted by $| I_n |$, the length $l\{I_n\}$ (in which the I_n need not be disjoint) will be considered to be the sum of the series $\Sigma \, | I_n |$, provided this series is convergent. The sequence $\{I_n\}$ is said to cover a set D of real numbers if any number in D is contained in at least one of the intervals I_n. A set D is said to be a *0-set* if there exists, for any given $\epsilon > 0$, a sequence $\{I_n\}$ covering D such that $l\{I_n\} < \epsilon$. For the easier part of the fundamental theorem use is made of the following well known

[1] Cf. H. Lebesgue, *Leçons sur l'intégration*, (2nd ed., 1928), chap. II; E. Kamke, *Das Lebesguesche Integral* (1925), § 5; E. W. Hobson, *The Theory of Functions of a Real Variable*, (2nd ed., 1921), chap. VI.

LEMMA. *If $D_1, D_2, \cdots$ is a sequence of 0-sets and D is the set of those numbers contained in at least one D_n, then D is a 0-set.*

In fact, if $\epsilon > 0$ is given, $\{I_n^m\}$ is a sequence of open intervals covering D_m, and $l\{I_n^m\} < \epsilon/2^m$, then the enumerable collection of open intervals in all $\{I_n^m\}$ covers D and its length is less than ϵ.

4. Let $f(x)$ be a real-valued function defined on the interval $a \leqq x \leqq b$, which will be denoted by J. If $f(x)$ is bounded on J, the *oscillation* of $f(x)$ on an interval K, which has at least one point in common with J, is defined as the difference between the least upper bound and the greatest lower bound of $f(x)$ on K, while the *fluctuation* of $f(x)$ at a given point x of J is the greatest lower bound of the oscillations of $f(x)$ on all intervals having x as interior point. Thus $f(x)$ is continuous at x if and only if its fluctuation at x is zero.

5. Let Δ (defined by x_i, $i = 0, \cdots, n$) be a partition of J and let Γ (defined by y_i, $i = 1, \cdots, n$) be an intermediate partition of Δ. The Riemann sum $S(\Delta, \Gamma)$ of $f(x)$ is defined as the sum

$$S(\Delta, \Gamma) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(y_i),$$

and the function $f(x)$ is said to have a Riemann integral if there exists a number S such that one can choose for every $\epsilon > 0$ a $\delta > 0$ which has the property that

$$| S - S(\Delta, \Gamma) | < \delta, \quad \text{whenever} \quad d(\Delta) < \epsilon.$$

FUNDAMENTAL THEOREM. *A function $f(x)$ defined on J has there a Riemann integral if and only if $f(x)$ is bounded on J and the set D of all discontinuity points of $f(x)$ is a 0-set.*

The necessity of these conditions will be proved in 6, their sufficiency in 7, 8, 9.

6. If $f(x)$ is not bounded, it is not bounded on at least one segment of any partition of J. Hence $S(\Delta, \Gamma)$ is, for any fixed Δ, an unbounded function of Γ, and so $f(x)$ cannot have a Riemann integral.

Now let $f(x)$ be bounded. Let the set of points at which the fluctuation of $f(x)$ is more than $1/n$ be denoted by D_n. Then the set D of all discontinuity points of $f(x)$ is the set of points contained in at least one D_n. If D is not a

0-set, it follows from the lemma in 3 that there exists an integer k such that D_k is not a 0-set. Hence there exists a $\gamma > 0$ such that any sequence of open intervals covering D_k has a length larger than γ. But then, if Δ is any partition of J, the total length of those segments of Δ which contain points of D_k is more than γ. Since the oscillation of $f(x)$ in these intervals is more than $1/k$, the oscillation of $S(\Delta, \Gamma)$ as a function of Γ is more than γ/k for any fixed Δ. Hence $f(x)$ cannot have a Riemann integral.

7. Let $f(x)$ be bounded on J, say $|f(x)| < c$, and let the set D of its discontinuity points be a 0-set. Let $\epsilon > 0$ be given, let $\{I_n\}$ be a sequence of open intervals covering D and such that $l\{I_n\} < \epsilon$, and let $\{K_n\}$ be the collection of those open intervals with rational end points on which the oscillation of $f(x)$ is less than ϵ. Since $\{I_n\} + \{K_n\}$ obviously covers J, a finite set covering J can be selected from $\{I_n\} + \{K_n\}$ in view of the Borel covering theorem. This finite set $I_1, \cdots, I_p$; $K_1, \cdots, K_q$ has then the following properties: The set covers J, the oscillation of $f(x)$ on each of the intervals $K_1, \cdots, K_q$ is less than ϵ, finally $|I_1| + \cdots + |I_p| < \epsilon$. Let δ be a positive number less than half the distance of any two distinct endpoints of the intervals $I_1, \cdots, I_p$; $K_1, \cdots, K_q$.

8. Let Δ, Δ' be two partitions of J such that $d(\Delta) < \delta$, $d(\Delta') < \delta$ and let Γ, Γ' be arbitrary intermediate partitions of Δ, Δ'. Let the distinct points occurring in Δ and Δ' together be arranged in increasing order and denoted by $x_0 = a, x_1, \cdots, x_n = b$. Then

$$(*) \qquad S(\Delta, \Gamma) - S(\Delta', \Gamma') = \sum_{i=1}^{n} (x_i - x_{i-1})\{f(y_i) - f(y'_i)\},$$

where y_i, y'_i are those points of Γ, Γ' contained in the segments s, s' of Δ, Δ' respectively, and s and s' both contain the set $x_{i-1} \leqq x \leqq x_i$. Since the total length of s and s' together is less than 2δ, it follows from the definition of δ that s and s' both are contained either in *one* of the intervals $I_1, \cdots, I_p$ or in *one* of the intervals $K_1, \cdots, K_q$. Let Σ_1 and Σ_2 denote the sums of those terms on the right side of (*) for which the first and the second of these possibilities occur. The total length of the segments in Σ_1 is less than ϵ, since these segments do not overlap and are all contained in one of the intervals $I_1, \cdots, I_p$. Since $|f(x)| < c$, hence $|f(y_i) - f(y'_i)| < 2c$ it follows that $|\Sigma_1| < 2c\epsilon$. The total length of the segments in Σ_2 is not larger than the length $b - a$ of J. Since for each of the latter segments $|f(y_i) - f(y'_i)| < \epsilon$ in view of the fact that the oscillation of $f(x)$ on any interval J_m is less than ϵ, it follows that $|\Sigma_2| < (b - a)\epsilon$. Thus

$$(^{**}) \quad | S(\Delta, \Gamma) - S(\Delta', \Gamma') | \leqq | \Sigma_1 | + | \Sigma_2 | < (b - a + 2c)\epsilon, \text{ if } d(\Delta) < \delta,\ d(\Delta') < \delta.$$

9. Since $| S(\Delta, \Gamma) | < c(b - a)$, it is possible to select a sequence of partitions Δ_n, Γ_n in such a way that $d(\Delta_n) \to 0$ and that the sequence $S(\Delta_n, \Gamma_n)$ has a limit. On denoting this limit by S and substituting into $(^{**})$ for Δ', Γ' successively the partitions Δ_n, Γ_n, it follows by passing to the limit that

$$| S - S(\Delta, \Gamma) | \leqq (b - a + 2c)\epsilon, \text{ whenever } d(\Delta) < \delta.$$

This completes the proof of the fundamental theorem.

10. The method in 8 can be used to prove in a simple manner the following fact: If $f(x)$ is defined on J, and $\{I_n\}$ is a sequence of mutually disjoint intervals covering the set D of discontinuity points of $f(x)$, finally O_n is the oscillation of $f(x)$ on I_n if $f(x)$ is bounded on I_n, then O_n is defined for sufficiently large n and $\lim_{n \to \infty} O_n = 0$.

ADDENDUM: After correcting the proofs I noticed that a paper of Professor A. B. Brown, appearing in the last issue of the *American Mathematical Monthly* (vol. 43, 1936, pp. 396-398), contains a treatment of the same subject along lines very similar to those of the present note.

THE JOHNS HOPKINS UNIVERSITY.

ON THE CANONICAL TRANSFORMATIONS OF HAMILTONIAN SYSTEMS.

By E. R. VAN KAMPEN and AUREL WINTNER.

1. *Comparison of the method of the present paper with the usual treatment.* The object of this paper is to present an approach to the theory of canonical transformations. This approach is, in contrast to the usual treatment, one based directly on point transformations of a phase-space into another phase-space and is, therefore, symmetric in the coördinates and impulses. By this approach one can proceed in a manner which is quite straightforward and, in the main, algebraical in nature. In fact, the method in question is the extension to the case of a general canonical transformation of the algebraic treatment previously given by one of the authors [1] for the case of linear conservative canonical transformations. The difference between this particular case and the general case is the same as that between tensor algebra and tensor analysis. Correspondingly, it would be desirable to adapt the local theory of displacements in differential geometry to the group of canonical transformations, conservative or not. There also arises the question whether or not it is possible to imbed Birkhoff's theory of successive normalizations [2] into the tensor-analytical treatment. This question is connected with the problem whether or not it is possible to obtain Birkhoff's two invariants [3] of a non-degenerate fixed-point of an area-preserving surface transformation by an adaptation to canonical transformations of the treatment by means of which the partly topological invariant of the total curvature is obtained in the theory of surfaces by tensor-analytical methods, thus eliminating the use of formal series.

It turns out that the Jacobian matrices of canonical transformations define linear substitutions under which the invariant bilinear form of the complex group is *relative* invariant,[4] and that the multiplier, which will be denoted by s, is independent both of the position in the $2n$-dimensional phase-space and of the time t. In the usual theory, based on Pfaffians, the Jacobian

[1] A. Wintner, "On the linear conservative dynamical systems," *Annali di Matematica*, ser. 4, vol. 13 (1934), pp. 105-112.

[2] G. D. Birkhoff, *Dynamical Systems*, New York (1927), Chap. III.

[3] G. D. Birkhoff, *loc. cit.*, Chap. VIII.

[4] A. Wintner, *loc. cit.*, p. 107.

matrices turn out to define linear substitutions under which the bilinear form of the complex group is an *absolute* invariant. Actually, the invariance is not absolute but relative. In fact, an analysis of the usual treatment [5] reveals a mistake in the latter, since from the fact that two Pfaffians vanish simultaneously one is not justified to conclude that the two Pfaffians are identical, but merely that they are proportional. Correspondingly, the condition $s = \text{const.}$, which in the theory to be presented is *proved* as the integrability condition of a system of partial differential equations, is in the traditional treatment implied by the tacit *hypothesis* $s = 1$. As a matter of fact, s may be any non-vanishing real constant. However, since s turns out to be a non-vanishing *constant*, one can reduce the case of an arbitrary s to the case $s = 1$ by means of a linear conservative transformation which alters the Hamiltonian functions.

The situation with regard to the relative invariance is illustrated by the following point: It is shown [6] in the usual theory by means of Lagrangian brackets that if a canonical transformation is "completely canonical," then the transformation is volume-preserving, the *square* of the Jacobian being 1. Actually, it will be seen without any calculation that if s is assumed to be 1, then, whether the canonical transformation is completely canonical or not, the transformation is always volume-preserving. Furthermore, it follows from a theorem of Frobenius [7] on relative invariants that if $s = 1$, then the Jacobian also is 1, so that the transformation not only is volume-preserving but *orientation-preserving* as well. This generalizes the fact, well-known from the theory of surface transformations in the particular case $n = 1$, that the area-preserving transformations considered there are all orientation-preserving.[8]

2. *Hamiltonian matrices.* For a fixed n, let G denote the $2n$-rowed square matrix

$$(1) \qquad G = \begin{pmatrix} \omega & -\epsilon \\ \epsilon & \omega \end{pmatrix},$$

where ω denotes the n-rowed zero matrix and ϵ the n-rowed unit matrix. Thus

---

[5] Cf., e. g., G. Prange, "Die allgemeinen Integrationsmethoden der analytischen Mechanik," *Encyklopädie der Mathematischen Wissenschaften*, vol. IV, 1_{II} (1935), p. 757.

[6] Cf. G. Prange, *loc. cit.*, pp. 768-769.

[7] G. Frobenius, "Ueber die schiefe Invariante einer bilinearen oder quadratischen Form," *Journal für reine und angewandte Mathematik*, vol. 86 (1876), pp. 44-71, more particularly p. 48.

[8] G. D. Birkhoff, *loc. cit.*, Chap. VIII. Cf. also T. Levi-Civita and U. Amaldi, *Lezioni di Meccanica Razionale*, vol. II_2, Bologna (1927), p. 318.

(2) $$G = -G^{-1} = -G',$$

where the prime denotes the operation of transposition. Obviously, G is the matrix of what is called in the theory of Pfaffians the bilinear covariant and in algebra the invariant of the usual representation of the complex group, i. e., G is the normal form of an arbitrary non-singular skew-symmetric matrix.

A $2n$-rowed square matrix C will be said to be a Hamiltonian matrix if it is real and satisfies the condition

(3) $$C'GC = sG$$

for some non-vanishing number s, which will be called the multiplier of C. Since $\det G \neq 0$, it is clear from (3) that

(4) $$|\det C| = |s|^n,$$

so that, since $s \neq 0$, the Hamiltonian matrices are non-singular. It is easily verified from (3) that the Hamiltonian matrices form a group and that the multiplier of the Hamiltonian matrix $C^{(2)}C^{(1)}$ is $s^{(2)}s^{(1)}$, if $s^{(i)}$ is the multiplier of the Hamiltonian matrix $C^{(i)}$, where $i = 1, 2$. Since G is a Hamiltonian matrix in view of (2), and since (3) may be written in the form $C' = sGC^{-1}G^{-1}$, it follows that if C is a Hamiltonian matrix, then so is C', and that C and C' have the same multiplier. It is also seen that if $s = 1$, then C and C^{-1} have the same characteristic numbers and the characteristic equation of C is a reciprocal equation.

For an arbitrary non-vanishing real number c, put

(5) $$C_1 = \begin{pmatrix} c\epsilon & \omega \\ \omega & c\epsilon \end{pmatrix}, \qquad C_2 = \begin{pmatrix} c\epsilon & \omega \\ \omega & -c\epsilon \end{pmatrix}.$$

It is easily verified that these two matrices are Hamiltonian matrices and that if s_1, s_2 denote their multipliers, then

(6_1) $$s_1 = c^2, \qquad \det C_1 = s_1{}^n;$$

(6_2) $$s_2 = -c^2, \qquad \det C_2 = s_2{}^n.$$

These examples show that there exist for every n Hamiltonian matrices which have an arbitrary non-vanishing real number as multiplier.

It follows that (4) may be replaced by the more precise relation

(7) $$\det C = s^n.$$

In fact, since the multiplier of $C^{(2)}C^{(1)}$ is the product of the multipliers of

$C^{(2)}$ and $C^{(1)}$, and since there exists for every n and for every $s \neq 0$ a Hamiltonian matrix, it is clear that (7) holds for an arbitrary s if it holds in the case $s = 1$. In other words, it is sufficient to know that if C is a matrix for which $C'GC = G$, then $\det C = 1$. Now that $C'GC = G$ implies $\det C = 1$ is easily seen from (1) in view of a theorem of Frobenius on relative invariants.[9]

3. *A lemma on gradients.* Let an m-rowed square matrix $A = A(Y)$ and a vector $B = B(Y)$ with m components be defined in an m-dimensional domain of the space of the vector $Y = (y_1, \cdots, y_m)$. Suppose that the pair $A(Y)$, $B(Y)$ has the property that one can find for *every* scalar function $f = f(Y)$ another scalar function $g = g(Y)$ such that

$$B + A \operatorname{grad} f = \operatorname{grad} g \tag{8}$$

holds in the whole Y-domain under consideration. All functions and variables are supposed to be real. Let it be assumed for simplicity that A, B and f, g have continuous partial derivatives of the first and second order respectively. It is understood that $A \operatorname{grad} f$ in (8) denotes a matrix product, while a vector is a matrix with m rows and 1 column.

Now there exists[10] for a given pair A, B a suitably chosen g for *every* f if and only if B is a gradient and A of the form sE, where s is a scalar independent of Y and E denotes the m-rowed unit matrix. The sufficiency of these conditions is obvious from (8). In order to prove their necessity, suppose that there exists a g for every f. On choosing $f = f(Y)$ independent of Y, it is clear from (8) that B is a gradient. Thus (8) implies that $A \operatorname{grad} f$ is a gradient for every f. Hence, on choosing $f = f(y_1, \cdots, y_m)$ as an arbitrary polynomial $p(y_i)$ of the i-th component of $Y = (y_1, \cdots, y_m)$ and denoting by $A_i = A_i(Y)$ the vector the components of which are the elements of the i-th column of $A = A(Y)$, it is seen that the vector $q(y_i)A_i(Y)$ is a gradient for every polynomial $q(y_i)$. It follows, therefore, from the integrability condition satisfied by a gradient that $A(Y)$ is a diagonal matrix and that the i-th diagonal element of $A(Y)$ is a function of y_i alone. Since $A \operatorname{grad} f$ is a gradient for every f, on choosing i and j arbitrarily and placing $f(Y) = y_i y_j$, it follows that all diagonal elements of $A(Y)$ are equal. This proves that $A = sE$ for some constant scalar s.

---

[9] As to § 2, cf. A. Wintner, *loc. cit.*, p. 107.

[10] This fact is the analytical analogue to the algebraical fact used by A. Wintner, *loc. cit.*, p. 106.

4. *Canonical transformations.* The prime being the symbol for the transposition, let total differentiation with respect to the time t be denoted by a dot. Partial differentiations with respect to t or a coördinate or an impulse will be denoted by subscripts, as illustrated by the relation $\dot{a}(q, t) = a_q \dot{q} + a_t$. Thus a canonical system with n degrees of freedom may be written in the form

$$(9) \qquad \dot{q}_i = h_{p_i}, \qquad \dot{p}_i = -h_{q_i}, \qquad (i = 1, \cdots, n),$$

where the Hamiltonian function $h = h(q_1, \cdots, q_n; p_1, \cdots, p_n; t)$ may contain t and will be supposed to have continuous partial derivatives of the second order in the $(2n+1)$-dimensional domain under consideration. Put

$$(10) \qquad x_i = q_i, \quad x_{i+n} = p_i, \qquad (i = 1, \cdots, n),$$

and let X denote the vector with the $2n$ components $x_1, \cdots, x_{2n}$. Then (9) may be written in view of (1) in the form

$$(11) \qquad G\dot{X} = \operatorname{grad}_X h, \quad \text{where} \quad h = h(X; t).$$

The subscript of grad refers to the space in which one carries out the partial differentiations.

Now transform the phase-space X for every fixed t under consideration into a phase-space Y by means of a transformation

$$(12) \quad Y = Y(X; t), \quad \text{where} \quad Y = (y_1, \cdots, y_{2n}), \quad X = (x_1, \cdots, x_{2n}).$$

Suppose that the $2n$ functions (of $2n + 1$ variables) which occur in these transformation formulae have continuous partial derivatives of the second order and that the transformation determines for every fixed t a locally one-to-one correspondence between the two phase-spaces, the $2n$-rowed Jacobian being everywhere distinct from zero. A transformation (12) which satisfies these requirements is said to be a canonical transformation if it transforms *every* canonical system (11) into a system of differential equations which is again canonical, i. e., of the form

$$(13) \qquad G\dot{Y} = \operatorname{grad}_Y h^*, \quad \text{where} \quad h^* = h^*(Y; t),$$

it being understood that the new Hamiltonian function h^* need not be the same function as the Hamiltonian function h of (11).

If a canonical transformation (12) is such that h^* always is the same function as h, i. e., if $h^*(Y(X; t); t) = h(X; t)$ holds for every h, and if in

addition t does not occur in (12), then (12) is said to be a completely canonical transformation.[11]

5. *Partial differential equations characterizing canonical transformations.* Let

$$X = X(Y; t) \tag{14}$$

be the inverse of the transformation (12). Let C denote the Jacobian matrix of the transformation (12) for a fixed t, so that the j-th element in the k-th column of the non-singular $2n$-rowed square matrix C is the partial derivative of y_j with respect to x_k. The matrix C may be considered in view of (14) as a function $C(Y; t)$ of Y and t instead of as a function of X and t. It is easily verified by straightforward differentiations that

$$\dot{Y} = C\dot{X} + Y_t \tag{15}$$

and that, for an arbitrary f,

$$\operatorname{grad}_X f = C' \operatorname{grad}_Y f, \tag{16}$$

where C' is the transposed matrix, finally that

$$\operatorname{grad}_Y Y_t = C_t C^{-1}, \tag{17}$$

where A_t denotes the matrix or vector obtained by partial differentiation of every element of A, and the gradient of a vector denotes, of course, a matrix. Needless to say, Y_t is understood in the sense that one first differentiates (12) partially with respect to t and expresses then X by means of (14) as a function of Y and t. It is not assumed that (12) is a canonical transformation.

It is clear from (15) and (16) that (12) transforms (11) into

$$GC^{-1}(\dot{Y} - Y_t) = C \operatorname{grad}_Y h,$$

a relation which may be written in view of (1) in the form

$$G\dot{Y} = GY_t - GCGC' \operatorname{grad}_Y h.$$

On comparing this with (13), it is seen that (12) is a canonical transformation if and only if there exists for every h an $h^* = h^*(Y; t)$ such that

$$GY_t - GCGC' \operatorname{grad}_Y h = \operatorname{grad}_Y h^*, \tag{18}$$

where the Hamiltonian function $h = h(X; t)$ of (11) is thought of as expressed by means of (14) as a function of Y and t. On keeping t fixed and

[11] Cf. T. Levi-Civita and U. Amaldi, *loc. cit.*, p. 314.

comparing then (18) with (8), it follows from the lemma of § 3 that (12) is a canonical transformation if and only if the vector $B = GY_t$ is, for every fixed t, the gradient of a scalar function with respect to Y and the matrix $A = -GCGC'$ is, for every fixed t, of the form sE, where s is a scalar independent of Y and E denotes the $2n$-rowed unit matrix. This means in view of (1) that the transformation (12) is canonical if and only if there exist two scalars

$$(19_1) \qquad r = r(Y;t); \qquad (19_2) \qquad s = s(t).$$

such that

$$(20_1) \qquad GY_t = \text{grad}_Y r; \qquad (20_2) \qquad C'GC = sG.$$

If this condition is satisfied, the Hamiltonian function of the system (13) into which (11) is transformed by (12) is

$$(21) \qquad h^* = sh + r$$

in view of (18).

6. *The integrability conditions of* (20_1). Condition (20_1) cannot be applied directly to a given transformation (12), since the function r is unknown. It is, however, easy to see that there exists, for a given transformation (12), a function (19_1) satisfying (20_1) if and only if the matrix $C'GC$, considered as a function of Y and t, is independent of t.

In fact, there exists, for a given transformation (12), an r satisfying (20_1) if and only if $\text{grad}_Y GY_t$ is a symmetric matrix at every point of the $(2n+1)$-dimensional region under consideration. Now it is easily verified that $\text{grad}_Y GY_t = G\,\text{grad}_Y Y_t$, so that $\text{grad}_Y GY_t = GC_tC^{-1}$ in view of the identity (17). Hence there exists an r if and only if GC_tC^{-1} is a symmetric matrix, i. e., equal to the matrix $(GC_tC^{-1})'$. Now this condition means in view of (2) that

$$C'_tGC + C'GC_t \equiv (C'GC)_t$$

is the zero matrix for every Y and t, i. e., that $C'GC$ is a function of Y alone.

7. *Characterization of the canonical transformation by means of Hamiltonian matrices.* According to (19_2), the scalar s occurring in the necessary and sufficient conditions (20_1), (20_2) is, for every fixed t, independent of the position Y in the phase-space. This was a consequence of the lemma of § 3. Now it is easy to see that the pair of conditions (20_1), (20_2) implies that the scalar s must be independent not only of Y but of t as well. In fact, it will be shown that *a transformation* (12) *is a canonical transformation if and only*

if there exists an $s \neq 0$ which is independent both of t and Y and is such that the Jacobian matrix $C = C(Y; t)$ of (12) is, for every Y and t, a Hamiltonian matrix which has, in the sense of § 2, the constant s as multiplier.[12]

In fact, (12) is canonical if and only if (20_1) and (20_2) can be satisfied by suitably chosen functions (19_1) and (19_2). Now (20_1) is, according to § 6, satisfied by some function (19_1) if and only if $C'GC$ is independent of t, which means in view of (20_2) and (19_2) that s is a constant. This clearly proves the statement, since, the Jacobian being non-singular, $s \neq 0$ in view of (20_2).

The criterion thus proved may be formulated as follows: A transformation (12) is a canonical transformation if and only if its Jacobian matrix $C = C(Y; t)$ determines a linear substitution under which the bilinear form in cogradiant variables which belongs to the skew-symmetric matrix (1) is relative invariant with a multiplier $s \neq 0$ which is independent both of Y and t.

8. *The distortion of a canonical transformation.* If one calls the determinant of $C = C(Y; t)$ the distortion of (12), it is seen from (7) and from the integrability condition $s = \text{const.}$ found in § 6 that the distortion of a canonical transformation is independent both of Y and t.

It is also seen that if $s = 1$, then the distortion is 1, so that a canonical transformation with the multiplier $s = 1$ not only is volume-preserving, as shown by (4), but orientation-preserving as well, as shown by (7).

These results do not assume that t does not occur in (12) and are, therefore, more general than when formulated for the particular case of completely canonical transformations. In fact, $s = 1$ for every completely canonical transformation, while $s = 1$ does not imply that the transformation is completely canonical. This is clear from the relation (21) which shows that a canonical transformation is completely canonical if and only if the multiplier s is 1 and the function (19_1) is independent of Y.

Incidentally, the result $\det C = 1$ is new in the completely canonical subcase of the case $s = 1$ also, since the usual treatment of the completely canonical case gives[13] $\det C = \pm 1$. The proof of this weaker result is usually based on an application of the Lagrangian brackets or, what is the same thing, on the reciprocal matrix of the brackets of Poisson.

---

[12] Since the invariant matrix G is real and skew-symmetric, the matrix iG is Hermitian. Thus one might attempt to treat the question of relative invariance with respect to G by starting with the group under which the fundamental surface of an Hermitian space is invariant. This approach would necessitate discussions of reality which are analogous to those treated in the calculus of spinors.

[13] Cf. G. Prange, *loc. cit.*, p. 769.

The usual characterization [14] of the completely canonical transformation by means of the brackets of Lagrange or Poisson follows from (20_2) and (1) without any calculation, since $s = 1$ in the completely canonical case. Furthermore, it is not necessary to assume that the transformation is completely canonical but merely that it is canonical and has the multiplier $s = 1$.

As pointed out in § 1, the usual treatment starts with the implicit *assumption* that $s = 1$, an assumption which implies, by (7), that $\det C = 1$.

9. *The composition rule of canonical transformations.* It is clear from the definition of a canonical transformation that the set of canonical transformations defined in a common $(2n+1)$-dimensional region forms a group. The composition rule of the Jacobian matrices and of the multipliers of canonical transformations is clear from (20_2) and § 2. As far as the composition rule of the function (19_1) is concerned, it is easily verified that if $r_1 = r_1(Y; t)$ belongs to a canonical transformation which is substituted into another canonical transformation, and if $r_2 = r_2(Y; t)$ and s_2 belong to the latter transformation, then the function r belonging to the composite transformation is $r = s_2 r_1 + r_2$. In particular, if C, r, s belong to a canonical transformation, then C^{-1}, $-r/s$, $1/s$ belong to the inverse transformation.

Needless to say, the multiplier s is, according to (20_2), uniquely determined by the canonical transformation, while the function (19_1) remains, according to (20_1), undetermined up to an additive term which is a function of t alone.

On uniting a canonical transformation of a $2n_1$-dimensional phase-space with a canonical transformation of a $2n_2$-dimensional phase-space, it is often stated that one obtains a canonical transformation of a $(2n_1 + 2n_2)$-dimensional phase-space. Actually, while this is true under the traditional hypothesis $s = 1$, it is false in the general case. In fact, it is clear from the above rules that the united transformation is canonical if and only if $s_1 = s_2$, where s_1 and s_2 denote the multipliers of the two partial transformations.

10. *Another form of the criterion of* § 7. Let it here be mentioned for later application that if (12) satisfies (20_1) for every Y and t and (20_2) for a fixed t and every Y, then (12) is a canonical transformation. In fact, if (20_1) is satisfied for every Y and t, then $C'GC$ is, according to § 6, independent of t, so that (20_2) holds for every Y and t, if it holds for every Y and for a single t.

In order to formulate a converse of the fact thus proved, let a trans-

[14] Cf. G. Prange, *loc. cit.*, pp. 768-769.

formation (12) be termed a conservative transformation if it does not contain t. Now if (12) is a canonical transformation and t_0 denotes any fixed value of t, then the conservative transformation defined by $Y = Y(X; t_0)$ is a canonical transformation. In fact, any conservative transformation satisfies (20_1) by choosing $r \equiv 0$, and it satisfies (20_2) either for every t or for no t.

It may be mentioned that (12) need not be a canonical transformation even if the conservative transformation $Y = Y(X; t_0)$ is a canonical transformation for *every* fixed $t \doteq t_0$. This is clear from the criterion of § 7.

11. *An existence and uniqueness theorem for the initial value problem of canonical transformations.* So far it has been assumed that there is given a transformation (12) and the question was whether or not there exist a function $r = r(Y; t)$ and a constant $s \neq 0$ such that (12) satisfies (20_1) and (20_2). In what follows, the converse problem will be discussed. It will be shown that if there are given a fixed value t_0 of t, a conservative canonical transformation $Y = Y_0(X)$, finally a function $r = r(Y; t)$ which has continuous partial derivatives of the third order, then there exists exactly one canonical transformation (12) which reduces for $t = t_0$ to the given transformation $Y = Y_0(X)$ and belongs to the given scalar function $r = r(Y; t)$. According to § 7, the multiplier of the solution $Y = Y(X; t)$ of this problem is the multiplier of the given initial transformation. It is understood that t is to be restricted to a sufficiently small vicinity of t_0.

The proof proceeds as follows. First, one can write (20_1) with the use of (2) in the form

$$Y_t = - G \operatorname{grad}_Y r(Y; t), \tag{22}$$

where $r = r(Y; t)$ is the given function and $Y = Y(X; t)$ the unknown. Now consider (22) as a system (13) of ordinary differential equations with the independent variable t and assign for these ordinary differential equations the initial values $Y(X; t_0) = Y_0(X)$, where the point X of the phase-space is fixed. Let $Y = Y(X; t)$ be the corresponding solution of (22). According to the existence theorem of ordinary differential equations depending on parameters, the solution $Y = Y(X; t)$ exists in a portion of the space of the initial values X and in a sufficiently small interval $t_0 - b < t < t_0 + b$, where b can be chosen independent of X. Furthermore, the conditions of regularity made with regard to the given functions r, Y_0 imply that the solution $Y = Y(X; t)$ has continuous partial derivatives of the second order. Since the Jacobian of $Y(X; t)$ with respect to X is at $t = t_0$ the Jacobian of $Y_0(X)$, hence distinct from zero, it is clear from reasons of continuity that the Jacobian of $Y(X; t)$ does not vanish, if t is near enough to t_0. Thus (12) defines a

transformation. Now this transformation is, in view of § 10, a canonical transformation.

12. *Canonical integration constants.* Let $Y = Y(X; t)$ be the general solution of the canonical system (13), the $2n$ components of X being the $2n$ integration constants. Let it be assumed that these integration constants are combinations of initial values in such a way that the general solution $Y = Y(X; t)$ has continuous partial derivatives of the second order and that the Jacobian is nowhere zero. The integration constants represented by X are said to be canonical integration constants if the transformation (12) defined by means of the general solution of (13) is a canonical transformation.

Now the integration constants X are canonical integration constants if and only if the conservative transformation $Y = Y(X; t_0)$, where t_0 is one particular value of t, is a canonical transformation. This is clear without any calculation from the criterion of § 10, since (20_1) may be written in the form (22), while (22) is the same thing as (13), the vector X being a vector of integration constants.

It follows, in particular, that if $Y = Y(X; t)$ is the general solution of (13) and the integration constants represented by X are chosen to be the initial values of Y which belong to $t = t_0$, then (12) is a canonical transformation with the multiplier $s = 1$. In fact, (12) reduces then for $t = t_0$ to the identical transformation, hence to a canonical transformation with the multiplier $s = 1$.

If $Y = Y(X; t)$ is the general solution of

$$G\dot{Y} = \operatorname{grad}_Y \bar{h} \tag{23}$$

with canonical integration constants X, then (12) transforms every canonical system (11) into a canonical system (13), where

$$h^* = sh + \bar{h}, \tag{24}$$

s being the multiplier of (12). This is clear from (21), since $r = \bar{h}$ in view of (23) and (20_1). The rule (24) only means (cf. § 9) that the inverse transformation (14) transforms the Hamiltonian function $\bar{h}$ into the Hamiltonian function which is identically zero.

Remark. It may be mentioned that *every* canonical transformation (12) may be considered as one defined by means of some canonical integration constants X of a suitably chosen canonical system (13). This is clear from § 10, if one writes (20_1) in the form (13).

13. *Verification of the standard Pfaffian criterion.* Corresponding to (10), let in (12)

$$y_i = v_i, \qquad y_{i+n} = u_i, \qquad (i = 1, \cdots, n). \tag{25}$$

Then the usual criterion for a canonical transformation,[15] when adjusted to the fact that the constant $s \neq 0$ need not be 1 (cf. § 1), may be expressed by saying that (12) is a canonical transformation if and only if there exists a function r for which the Pfaffian

$$s \sum_{i=1}^{n} p_i dq_i - \sum_{i=1}^{n} u_i dv_i + r dt \tag{26}$$

becomes a complete differential in virtue of (12). This Pfaffian condition contains differentials of some of the $2n$ old variables and of some of the $2n$ new variables and is therefore, in contrast with the criterion of § 7, an unsymmetric criterion. On the other hand, it is easy to deduce this criterion from the theory developed above. In the verification use will be made of the fact that the alternating derivative (Stokesian) of a covariant vector is a tensor.

For a given transformation (12) which need not be canonical, consider the Pfaffian

$$\tfrac{1}{2} sX'GdX - \tfrac{1}{2} Y'GdY + r dt, \tag{27}$$

where $Z'AW$ denotes the bilinear form belonging to the square matrix A and to the vectors Z, W, so that $Z'AW$ is considered in the usual manner as a product of three matrices. It is easily verified from (1) that, whether the transformation (12) is canonical or not, the difference of the two Pfaffians (26), (27) is, in virtue of (1), (10) and (25), a complete differential, namely that of the function

$$s \sum_{i=1}^{n} p_i q_i - \sum_{i=1}^{n} u_i v_i.$$

Now (27) is a complete differential in virtue of (12) if and only if (12) is a canonical transformation. In fact, (27) becomes in virtue of (12) the Pfaffian

$$\tfrac{1}{2}(sX'G - Y'GC) dX + (r - \tfrac{1}{2} Y'GY_t) dt, \tag{28}$$

where Y and Y_t are considered as functions of X and t. It is clear that the Pfaffian (28) is a complete differential of a function of $2n + 1$ variables represented by X and t if and only if (i) the partial derivative of the covariant vector

[15] G. Prange, *loc. cit.*, p. 758.

(29) $$\tfrac{1}{2}(sX'G - Y'GC)$$

with respect to t is the (covariant) gradient of the scalar

(30) $$r - \tfrac{1}{2}Y'GY_t$$

with respect to the vector X and (ii) the alternating derivative of the covariant vector (29) with respect to the vector X is identically zero. Now the conditions (i), (ii) are equivalent to the conditions (20_1), (20_2) respectively.

In fact, condition (i) is satisfied if and only if

$$-\tfrac{1}{2}(Y'GC)_t = (\mathrm{grad}_X r)' - \tfrac{1}{2}[\mathrm{grad}_X(Y'GY_t)]',$$

a relation which, when transposed, may be written as

$$-\tfrac{1}{2}(C'G'Y)_t = \mathrm{grad}_X r - \tfrac{1}{2}\,\mathrm{grad}_X(Y'GY_t),$$

or, on using (1), (16) and the definition of C, also as

$$\tfrac{1}{2}C'_tGY + \tfrac{1}{2}C'GY_t = C'\,\mathrm{grad}_Y r - \tfrac{1}{2}C'GY_t + \tfrac{1}{2}C'_tGY,$$

i. e., as

$$C'GY_t = C'\,\mathrm{grad}_Y r.$$

Since this may be written in the form (20_1), it follows that condition (i) is equivalent to (20_1).

On the other hand, since an alternating derivative is transformed by any transformation (12) as a tensor, and since the alternating derivative of $X'G$ with respect to X is G in view of (1), it follows that the alternating derivative of $Y'GC$ with respect to X is $C'GC$ and that condition (ii) is satisfied if and only if the matrix

$$\tfrac{1}{2}(sG - C'GC)$$

is identically zero. Hence condition (ii) is equivalent to (20_2).

Remark. While the above considerations were based on the transformations of a $2n$-dimensional phase-space which depend on t, it is possible to develop the above theory in a $(2n+2)$-dimensional space which is obtained by adjoining to the $2n$-dimensional phase-space the function (19_1) and the time t as a pair of canonically conjugate variables. This approach is particularly convenient when the method is applied to the reduction of the degree of freedom by means of canonical transformations based on known first integrals in involution.

THE JOHNS HOPKINS UNIVERSITY.

NOTE ON A CERTAIN BILINEAR FORM THAT OCCURS IN STATISTICS.[1]

By Allen T. Craig.

It is the purpose of this note to present some of the properties of the distribution of a real symmetric bilinear form of $2n$ normally but independently distributed variables.

Let x any y be independently distributed in accord with

$$f(x) = (1/\sqrt{2\pi}) \exp[-(x^2/2)] \quad \text{and} \quad g(y) = (1/\sqrt{2\pi}) \exp[-(y^2/2)].$$

Let $x_1, x_2, \cdots, x_n$ and $y_1, y_2, \cdots, y_n$ be n random pairs of values of x and y and let $A = \| a_{uv} \|$ $(u, v = 1, 2, \cdots, n)$, be a real symmetric matrix of rank r. If B denote the bilinear form $\sum_{u,v=1}^{n} a_{uv} x_u y_v$, the characteristic function $\phi(t)$ of the distribution of B is given by

$$\phi(t) = \left(\frac{1}{2\pi}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(itB - \tfrac{1}{2}Q_1 - \tfrac{1}{2}Q_2)\, dy_n dx_n \cdots dy_1 dx_1$$

where $Q_1 = \sum_1^n x_u^2$ and $Q_2 = \sum_1^n y_u^2$. Since A is a real symmetric matrix, there exists a real orthogonal matrix $L = \| l_{uv} \|$ such that

$$L'AL = \left\| \begin{array}{cccccc} \lambda_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & \cdots & \lambda_r & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{array} \right\|$$

where $\lambda_1, \cdots, \lambda_r$ are the r real, non-zero roots of the characteristic equation

[1] Presented to the American Mathematical Society on September 12, 1935.

of A, and L' is the conjugate of L. If we subject the x's and y's to the same linear homogeneous transformation with matrix L, then B becomes $\sum_1^r \lambda_u x'_u y'_u$ while Q_1 and Q_2 remain unchanged. Thus, upon dropping the primes, $\phi(t)$ becomes

$$\phi(t) = \left(\frac{1}{2\pi}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left(it \sum_1^r \lambda_u x_u y_u - \tfrac{1}{2}Q_1 - \tfrac{1}{2}Q_2\right) dy_n dx_n \cdots dy_1 dx_1$$
$$= [(1+\lambda_1^2 t^2)(1+\lambda_2^2 t^2)\cdots(1+\lambda_r^2 t^2)]^{-\frac{1}{2}}.$$

If ω_s represents the s-th semi-invariant of the distribution of B, it follows from

$$i^s \omega_s = \left.\frac{d^s \log \phi(t)}{dt^s}\right|_{t=0}$$

that

$$\omega_{2s+1} = 0 \quad \text{and} \quad \omega_{2s} = (2s-1)! \sum_1^r \lambda_u^{2s}.$$

In terms of the semi-invariants, the distribution function $F(B)$ of B is then given formally by

$$\exp\left[\sum_1^\infty \frac{(it)^s}{(s)!}\omega_s\right] = \int_{-\infty}^{\infty} \exp(itB) F(B)\, dB.$$

Thus,

$$F(B) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-itB + \sum_1^\infty [(it)^s/(s)!]\omega_s\} dt$$
$$= \psi(B) + A_2\psi^{(2)}(B) + A_4\psi^{(4)}(B) + \cdots,$$

where

$$\psi(B) = \frac{1}{\sigma\sqrt{2\pi}} \exp[-(B^2/2\sigma^2)], \qquad \sigma^2 = \sum_1^r \lambda_u^2, \qquad \psi^{(s)}(B) = \frac{d^s\psi(B)}{dB^s},$$

and the A's are the well known functions of the semi-invariants.[2]

Of particular interest is the case in which $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 1$ and the rank r is even, say $r = 2k$. For example, B may be n times the sample covariance,

$$B = \sum (x_j - \bar{x})(y_j - \bar{y}) = \sum_{u,v} a_{uv} x_u y_v$$

where $n\bar{x} = \Sigma x_u$, $n\bar{y} = \Sigma y_v$ and $a_{uv} = (n-1)/n$ or $-1/n$ according as $u = v$ or $u \neq v$ respectively. The n roots of the characteristic equation of the matrix $\| a_{uv} \|$ are $0, 1, 1, \cdots, 1$ and, if n is odd, the rank of the matrix is $r = n - 1 = 2k$. Under these conditions we have[3]

[2] Cf. T. N. Thiele, *Theory of Observations* (1903), pp. 33-35.

[3] Cf. J. Wishart and M. S. Bartlett, "The distribution of second order moment statistics in a normal system," *Proceedings of the Cambridge Philosophical Society*, vol. 28 (1931-32), p. 458.

$$\phi(t) = (1 + t^2)^{-k}$$

and

$$F(B) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\exp(-itB)}{(1+t^2)^k}\, dt.$$

By a suitably chosen contour for integration in the complex plane, it may be shown that

$$F(B) = \exp(-|B|)[a_0 + a_1|B| + a_2B^2 + a_3|B^3| + \cdots + a_{k-1}|B^{k-1}|]$$

where

$$a_s = \frac{(2k-s-2)!}{2^{2k-s-1}(k-s-1)!\,s!\,(k-1)!}, \qquad (s = 0, 1, \cdots, k-1).$$

It is interesting to observe [4] that this distribution function is identical with that of the sum of k independent values of x drawn at random from a population characterized by $f(x) = \frac{1}{2}\exp(-|x|)$.

THE UNIVERSITY OF IOWA.

[4] Cf. K. Mayr, "Wahrscheinlichkeitsfunktionen und ihre Anwendungen," *Monatshefte für Mathematik und Physik*, Bd. 30 (1920), p. 25.

THE APPLICATION OF THE THEORY OF ADMISSIBLE NUMBERS TO TIME SERIES WITH VARIABLE PROBABILITY.[1]

By Francis Regan.

I. *Introduction.* This paper is concerned with the consistency of the statistical theory of probability as applied to time series. In testing this consistency, we shall employ the theory of admissible numbers. A time series is a sequence of occurrences distributed independently in such a manner that there is a definite probability of an occurrence in any given interval of time. A time series may be represented by a sequence of points on the positive time axis. If E is any (Lebesgue measurable) set of points on this axis, then there is a definite probability $f(\alpha, E)$ that there will be α points of the time series in the set E. We shall assume that this function possesses a certain periodicity which is defined in terms of the following transformation. Let $T_y(x) = x + y$, i. e., the transformation $T_y(x)$ translates a point x into a point $x' = x + y$. This operation also transforms a set of points E into a set E' and this transformation will be denoted by the equation $E' = T_y(E)$. The function $f(\alpha, E)$ is said to possess a period A provided

$$f[\alpha, T_A(E)] = f(\alpha, E)$$

where A is independent of α. If all numbers A are periods of this function, the time series is said to possess a constant probability. Otherwise the probability will be said to be variable.

The mathematical idealization of a time series is an increasing sequence of positive numbers $s_1, s_2, \cdots$ whose properties are defined in terms of the expression $x(\alpha, E, A)$ where

$$x(\alpha, E, A) = x^{(1)}, x^{(2)}, x^{(3)}, \cdots, x^{(k)}, \cdots$$

$$x^{(k)} = \begin{cases} 1 & \text{if there are exactly } \alpha \text{ points of the series in } T_{(k-1)A}(E) \\ 0 & \text{otherwise.} \end{cases}$$

The set E is contained in the interval $0 < y \leq A$ and A is a period of $f(\alpha, E)$. Thus $x(\alpha, E, A)$ represents an event which succeeds on its k-th trial if and only if there are exactly α points of the time series belonging to the set

[1] This paper was presented to the Society June 22, 1933.

$T_{(k-1)A}(E)$. The success ratio for the first n trials of the event $x(\alpha, E, A)$ is denoted by $p_n[x(\alpha, E, A)]$ and is given by the equation

$$p_n[x(\alpha, E, A)] = \sum_{k=1}^{n} x^{(k)}/n.$$

The probability $p[x(\alpha, E, A)]$ of this event is the limit of the success ratio, i. e.,

$$p[x(\alpha, E, A)] = \lim_{n\to\infty} p_n[x(\alpha, E, A)].$$

Thus the time series must satisfy the condition $p[x(\alpha, E, A)] = f(\alpha, E)$. We shall make the further demand that $x(\alpha, E, A)$ is an element of the set $A[f(\alpha, E)]$, where $A[f(\alpha, E)]$ is the set of all admissible numbers associated with the probability $f(\alpha, E)$. We shall however restrict ourselves in Art. II to the consideration of sets E consisting of finite sums of intervals and to the function of variable probability while in § 4 of Art. III, we shall consider for constant probability sets E which are not of the form already discussed by the author.[2]

It is now clear that the assumption of periodicity of $f(\alpha, E)$ is necessary in order that the probability can have a meaning in terms of the statistical theory. It can be proved that

$$f(\alpha, E) = [K(E)]^{\alpha} e^{-K(E)}/\alpha!$$

where $K(E)$ is a non-negative absolutely additive, absolutely continuous function of point sets.[3]

II. *At least one point or exactly α points lying in an interval.*

1. In this section we shall concern ourselves with the restricted case in which E is the interval of length η, beginning at time s. The function $f(\alpha, E)$ shall be designated by $f(\alpha, \eta, s)$ and $K(E)$ shall be denoted by $K(\eta, s)$. The function $f(\alpha, \eta, s)$ is periodic in s, where the period depends upon the nature of the physical event; for example, let us assume that the probability of an

---

[2] See Regan, "The application of the theory of admissible numbers to time series with constant probability," *Transactions of the American Mathematical Society*, vol. 36 (July, 1934), pp. 511-529.

[3] Fry, *Probability and Its Engineering Uses*, pp. 216-227 and pp. 232-235, shows that under suitable restrictions the function $f(\alpha, E)$ must be expressible in the above form where $K(E)$ is a Riemann integral of a continuous function. In a paper written in conjunction with A. H. Copeland, entitled "A postutational treatment of the Poisson law," *Annals of Mathematics*, vol. 37, no. 2, April, 1936, pp. 357-362, the above equation under a more general set of hypotheses has been derived by the authors.

event happening at 9.00 a.m. Monday is the same as that at 9.00 a.m. Tuesday, etc. Here the primitive period is the day. Hence $f(\alpha, \eta, s)$ is periodic in s for a certain η. Let this period be m, then $f(\alpha, \eta, s) = f(\alpha, \eta, s + mn)$, $(n = 1, 2, \cdots)$. If this period is not one it may be made so by a linear transformation. Therefore there will be no loss in generality, if we confine our work to the case where the period is one.

We shall first consider the case of at least one point lying in an interval of length η, beginning at time s. A time series representing this phenomenon may be represented by a set of points $s_1 < s_2 < s_3 < \cdots < s_i < \cdots$ distributed along the positive s-axis. The probability of at least one point lying in any interval η of the time axis, beginning at time s, is $[1 - f(0, \eta, s)]$.

The set of points s_i may be obtained by a certain transformation applied to a time series $\tau_1, \tau_2, \tau_3, \cdots$ which has a constant probability, and which is such that if

$$\sim x_0(0, \tau, t, \Lambda) = x_0^{(1)}, x_0^{(2)}, \cdots, x_0^{(k)}, \cdots$$

where

$x_0^{(k)} = 1$ if there is at least one point τ_i in

$$I_k\colon\ t + (k-1)\Lambda < h \leq t + \tau + (k-1)\Lambda$$

otherwise

$$x_0^{(k)} = 0,$$

where $\Lambda = \delta \cdot \rho \cdot 2^{-\sigma+1}$, $t = \delta \cdot r \cdot 2^{-\sigma+1}$, $\tau = \delta \cdot m \cdot 2^{-\sigma+1}$; ρ, σ, r and m being integers such that $r + m \leq \rho$, and $\delta = K(1, 0)$, then $\sim x_0(0, \tau, t, \Lambda)$ is an element of $A[1 - e^{-\tau}]$.[4]

For our purpose here we will only consider those values of $\rho \cdot 2^{-\sigma+1}$ which are integers. Hence $\Lambda = \delta \cdot M$, where M is an integer.

It is now our problem to show how this series may be transformed into one which has a variable probability and which is consistent with the frequency theory.

Let $I_{\eta,s}$ be the interval $s < \xi \leq s + \eta$. Let $K(\eta, 0) = T(\eta) = \lambda$. It is assumed that $T(\eta + 1) = T(\eta) + \delta$. From the nature of the function $K(\eta, s)$, we know that the inverse $\eta = T^{-1}(\lambda)$ exists.[5] Then $T^{-1}(\lambda + \delta) = T^{-1}(\lambda) + 1$.

---

[4] A time series of this type has been constructed by the author in a paper entitled "The application of the theory of admissible numbers to time series with constant probability," *Transactions of the American Mathematical Society*, vol. 36 (July, 1934), no. 3, pp. 511-529. It should be noted that the constant factor δ that has been used here is nothing more than the k (page 512) which we chose as one, but for our purpose here k is chosen to be $\delta = K(1, 0)$.

[5] In the paper mentioned above written by Copeland and Regan, they show that if I_x is the interval $0 < y \leq x$ and $t = K(I_k) = T(x)$, then this transformation has a unique inverse $x = T^{-1}(t)$.

We shall now construct a time series $s_1 < s_2 < \cdots < s_i < \cdots$ such that $s_i = T^{-1}(\tau_i)$. Let

$$\sim x(0, \eta, s, M) = x^{(1)}, x^{(2)}, \cdots, x^{(k)}, \cdots.$$

where
$x^{(k)} = 1$ if there is at least one point in

$$I_k\colon s + (k-1)M < h \leq s + \eta + (k-1)M$$

otherwise

$$x^{(k)} = 0,$$

where $s = T^{-1}(t)$, $s + \eta = T^{-1}(t + \tau)$, t and τ being equal to $\delta \cdot r \cdot 2^{-\sigma+1}$ and $\delta \cdot m \cdot 2^{-\sigma+1}$ respectively. Since

$$T^{-1}[t + (k-1)\Lambda] = T^{-1}(t) + (k-1)M = s + (k-1)M$$

and

$$T^{-1}[t + \tau + (k-1)\Lambda] = T^{-1}(t+\tau) + (k-1)M = s + \eta + (k-1)M$$

it follows that $x^{(k)} = x_0{}^{(k)}$ and hence $\sim x(0, \eta, s, M)$ is an element of $A[1 - e^{-\tau}]$, but $t = T(s)$ and $t + \tau = T(s + \eta)$ or

$$\tau = T(s+\eta) - T(s) = K(\eta, s).$$

Therefore $\sim x(0, \eta, s, M)$ is an element of $A[1 - e^{-K(\eta, s)}]$ and

$$s_1 < s_2 < \cdots < s_i < \cdots$$

is a time series with the desired properties.

2. We shall investigate the case when within the interval I_k of length $T^{-1}(\tau = \delta \cdot m \cdot 2^{-\sigma+1})$, sub-intervals of the form $T^{-1}(\delta \cdot 2^{-\sigma+1})$ are omitted from consideration.

In the case of constant probability we have shown that if

$$\sim x(0, \tau', t, \Lambda) = x_1{}^{(1)}, x_1{}^{(2)}, \cdots, x_1{}^{(k)}, \cdots$$

where $x_1{}^{(k)}$ is one if there is at least one point in the n intervals of I_k, where the sub-intervals of I_k are:

$$t + t_i + (k-1)\Lambda < h \leq t + t_i + \delta \cdot 2^{-\sigma+1} + (k-1)\Lambda, \qquad (i = 1, 2, \cdots, n),$$

then $\sim x(0, \tau', t, \Lambda)$ is an element of $A[1 - e^{-\tau'}]$, where

$$\tau' = \tau'_1 + \tau'_2 + \cdots + \tau'_n, \qquad \tau'_i = \delta \cdot 2^{-\sigma+1},$$

beginning at time t_i. The numbers Λ, t, τ' and τ are $\delta \cdot M$, $\delta \cdot r \cdot 2^{-\sigma+1}$, $\delta \cdot n \cdot 2^{-\sigma+1}$ and $\delta \cdot m \cdot 2^{-\sigma+1}$ respectively where ρ, σ, r, M, m and n are integers with $n \leq m$ and the sub-intervals begin at $t_i = \delta \cdot \rho_i \cdot 2^{-\sigma+1}$, $(i = 1, 2, \cdots, n)$, where ρ_i is an integer, satisfying, $r \leq \rho_i \leq r + m - 1$.

Now we may use the same transformation as given in § 1. Let

$$\sim x(0, \eta', s, M) = x^{(1)}, x^{(2)}, \cdots, x^{(k)}, \cdots {}^{6}$$

where $x^{(k)}$ is one if there is at least one point in the n intervals of I_k, where the sub-intervals are:

$$s + t'_i + (k-1)M < h \leq s + t'_i + T^{-1}(\delta 2^{-\sigma+1}) + (k-1)M, \qquad (i = 1, 2, \cdots, n)$$

and zero otherwise, where $s = T^{-1}(t)$, $(t = \delta \cdot r \cdot 2^{-\sigma+1})$, $t'_i = T^{-1}(\delta \cdot \rho_i \cdot 2^{-\sigma+1})$, $s + \eta = T^{-1}(t + \tau)$, where $s \leq t'_i \leq s + \eta - T^{-1}(\delta \cdot 2^{-\sigma+1})$. Since

$$\begin{aligned} &T^{-1}[t + t_i + (k-1)\Lambda] \\ &\qquad = T^{-1}(t) + T^{-1}(t_i) + (k-1)M = s + t'_i + (k-1)M \end{aligned}$$

and

$$\begin{aligned} &T^{-1}[t + t_i + \delta \cdot 2^{-\sigma+1} + (k-1)\Lambda] \\ &\qquad = T^{-1}(t) + T^{-1}(t_i) + T^{-1}(\delta \cdot 2^{-\sigma+1}) + (k-1)M \\ &\qquad = s + t'_i + T^{-1}(\delta \cdot 2^{-\sigma+1}) + (k-1)M, \end{aligned}$$

it follows that $x^{(k)} = x_1{}^{(k)}$ and hence $\sim x(0, \eta', s, M)$ is an element of $A[1 - e^{-\tau'}]$, where $\tau' = \tau'_1 + \tau'_2 + \cdots + \tau'_n$, with each of the $\tau_i = \delta \cdot 2^{-\sigma+1}$ beginning at time t_i. Therefore, since $t_i = T(t'_i)$ and

$$t_i + \delta 2^{-\sigma+1} = T[t'_i + T^{-1}(\delta 2^{-\sigma+1})],$$

we have

$$\delta \cdot 2^{-\sigma+1} = T[t'_i + T^{-1}(\delta \cdot 2^{-\sigma+1}) - T(t'_i)] = K[T^{-1}(\delta 2^{-\sigma+1}), t'_i].$$

Therefore, $\sim x(0, \eta', s, M)$ is an element of

$$A[1 - \exp\{-\sum_{i=1}^{n} K[T^{-1}(\delta 2^{-\sigma+1}), t'_i]\}].$$

3. When we consider α points in an interval of length η, or the case

[6] The symbol η' represents the sum of the n sub-intervals $T^{-1}(\delta \cdot 2^{-\sigma+1})$ under consideration.

where there are α points in an interval where sub-intervals of the form $T^{-1}(\delta \cdot 2^{-\sigma+1})$ have been omitted from consideration, we need only apply §§ 1 and 2 to the corresponding time series of constant probability.

III. *Intervals not of the form* $(r+m)\cdot 2^{-\sigma+1}$.

4. We shall discuss for constant probability the case in which the interval becomes irrational or rational and not of the form $(r+m)\cdot 2^{-\sigma+1}$. It may be such that t and τ are both irrational, both rational and not of the forms $r\cdot 2^{-\sigma+1}$, $m\cdot 2^{-\sigma+1}$, or one may be rational and the other irrational.

Let us choose r_1, r_2, m_1, m_2, and Λ so that $r_1\cdot 2^{-\sigma+1} < t < r_2\cdot 2^{-\sigma+1}$, $(r_2+m_2)\cdot 2^{-\sigma+1} < t+\tau < (r_1+m_1)\cdot 2^{-\sigma+1}$, and $\Lambda=\rho\cdot 2^{-\sigma+1}$, where $r_1+m_1\leq\rho$ and $r_2+m_2\leq\rho$. Let $\tau_1=m_1\cdot 2^{-\sigma+1}$, $\tau_2=m_2\cdot 2^{-\sigma+1}$, $t_1=r_1\cdot 2^{-\sigma+1}$, $t_2=r_2\cdot 2^{-\sigma+1}$.

Given any positive number, ϵ, we can select r_1, r_2, ρ_1, ρ_2, σ such that

$$\text{(a)}\qquad \{1-f(0,\tau_1,t_1)\}^k-\epsilon/2\leq\{1-f(0,\tau,t)\}^k \leq\{1-f(0,\tau_2,t_2)\}^k+\epsilon/2.$$

It follows from Theorem 2 [7] that the numbers N_s can be chosen so that for every m, r, σ and Λ such that $m+r\leq\Lambda 2^{\sigma-1}$ that the number $\sim x(0,\tau_1,t_1,\Lambda)$ is a member of the set $A[1-f(0,\tau_1,t_1)]$. For the same numbers N_s, then the number $\sim x(0,\tau_2,t_2,\Lambda)$ is a member of the set $A[1-f(0,\tau_2,t_2)]$. Hence we can select a number N_0 such that

$$\text{(b)}\qquad \{1-f(0,\tau_2,t_2)\}^k-\epsilon/2\leq p_N[\prod_{i=1}^{k}(r_i/n)\sim x(0,\tau_2,t_2,\Lambda)\cdot],$$

and

$$\text{(c)}\qquad p_N[\prod_{i=1}^{k}(r_i/n)\sim x(0,\tau_1,t_1,\Lambda)\cdot]\leq\{1-f(0,\tau_1,t_1)\}^k+\epsilon/2$$

whenever $N>N_0$.

Now let us define a number $\sim x(0,\tau,t,\Lambda)$ such that its k-th digit is one if there exists at least one point of the time series in the interval I_k: $t+(k-1)\Lambda<h\leq t+\tau+(k-1)\Lambda$, where t and τ have been restricted as above and $\Lambda=\rho\cdot 2^{-\sigma+1}$. Then we have

$$\text{(d)}\qquad p_N[\prod_{i=1}^{k}(r_i/n)\sim x(0,\tau_2,t_2,\Lambda)\cdot]\leq p_N[\prod_{i=1}^{k}(r_i/n)\sim x(0,\tau,t,\Lambda)\cdot] \leq p_N[\prod_{i=1}^{k}(r_i/n)\sim x(0,\tau_1,t_1,\Lambda)\cdot].$$

[7] See author's memoir, *loc. cit.*, p. 522.

Subtracting ϵ from the second inequality of (a), we have

$$\{1-f(0,\tau,t)\}^k-\epsilon \leqq \{1-f(0,\tau_2,t_2)\}^k-\epsilon/2.$$

Combining with (b) we get

$$\text{(e)} \qquad \{1-f(0,\tau,t)\}^k-\epsilon \leqq p_N[\prod_{i=1}^{k}(r_i/n) \sim x(0,\tau_2,t_2,\Lambda)\cdot].$$

Adding ϵ to the first inequality of (a), we have

$$\{1-f(0,\tau_1,t_1)\}^k+\epsilon/2 \leqq \{1-f(0,\tau,t)\}^k+\epsilon.$$

Combining with (c) we get

$$\text{(f)} \qquad p_N[\prod_{i=1}^{k}(r_i/n) \sim x(0,\tau_1,t_1,\Lambda)\cdot] \leqq \{1-f(0,\tau,t)\}^k+\epsilon.$$

Using (e) and (f) with (d), we get

$$\begin{aligned}\{1-f(0,\tau,t)\}^k-\epsilon < p_N[\prod_{i=1}^{k}(r_i/n) \sim x(0,\tau,t,\Lambda)\cdot]\\ \leqq \{1-f(0,\tau,t)\}^k+\epsilon\end{aligned}$$

whenever $N > N_0$. Therefore, we have

$$p[\prod_{i=1}^{k}(r_i/n) \sim x(0,\tau,t,\Lambda)\cdot] = \{1-f(0,\tau,t)\}^k.$$

Hence, we see that for the same numbers N_s as found from Theorem 2, that $\sim x(0,\tau,t,\Lambda)$ is a member of the set $A[1-f(0,\tau,t)]$, where t and τ are not of the forms $r\cdot 2^{-\sigma+1}$ and $\rho\cdot 2^{-\sigma+1}$ respectively but $t+\tau < \Lambda$.

From the principles of admissible numbers, it follows that $x(0,\tau,t,\Lambda)$ is an element of the set $A[f(0,\tau,t)]$ for every t and τ, such that $t+\tau<\Lambda$.

We have now proved the following theorem:

THEOREM. *If the hypothesis* (H_1) *of Theorem 2 is satisfied, then for the same numbers* N_s *which may be obtained for this case of Theorem 2, it is true for every* t *and* τ, *such that* $t+\tau<\Lambda$, *that the corresponding number* $x(0,\tau,t,\Lambda)$ *is an element of* $A[f(0,\tau,t)]$.

By the transformation of § 1, we can show that the series for variable probability is consistent for intervals not of the form $(r+m)2^{-\sigma+1}$.

ST. LOUIS UNIVERSITY,
ST. LOUIS, MO.

CERTAIN RATIONAL r-DIMENSIONAL VARIETIES OF ORDER N IN HYPERSPACE WITH A RATIONAL PENCIL OF $(r-1)$-DIMENSIONAL VARIETIES OF LOWER ORDER.

By B. C. Wong.

1. *Introduction.* Let $M_{r-k}^{m^k}$ denote the $(r-k)$-dimensional manifold, of order m^k, of complete intersection of k given general hypersurfaces of order m in an r-space, S_r. The hypersurfaces of order n passing through $M_{r-k}^{m^k}$ form a linear system, $|V|$, of dimension[1]

$$(1)\quad \rho(r,n,m,k) = -1 + \binom{k}{1}\binom{n-m+r}{r} - \binom{k}{2}\binom{n-2m+r}{r} + \binom{k}{3}\binom{n-3m+r}{r} - \cdots - (-)^{k-1}\binom{k}{k-1}\binom{n-km+m+r}{r} - (-1)^k\binom{k}{k}\binom{n-km+r}{r},$$

where the second factor in any term is to be taken equal to zero if its upper number is less than r. It is implied that $n \geqq m$. If $n = m$, then $\rho(r,m,m,k) = k-1$ and we have the ∞^{k-1}-system determined by the k given hypersurfaces intersecting in $M_{r-k}^{m^k}$. We shall exclude this case from our consideration and shall, in what follows, assume $n > m$.

In this paper we are concerned with the rational r-dimensional variety, denoted by $\Phi_r^{N(r,n,m,k)}$ or just Φ_r, of order $N(r,n,m,k)$ in a $\rho(r,n,m,k)$-space, $\Sigma_{\rho(r,n,m,k)}$, which is representable upon S_r by the hypersurfaces of $|V|$. The results obtained will be, in some measure, generalizations of those already obtained by D. W. Babbage,[2] B. C. Wong,[3] and W. L. Edge.[4] We shall, in §§ 2, 3, 4, describe, in general terms, the properties of the representation and then, in §§ 5-8, describe, in some detail, the case $k = 2$, and, in particular, the case $n = m + 1 \geqq 3$. In conclusion, we shall mention, in § 9, some other special cases of interest.

2. *The determination of* $N(r,n,m,k)$. To determine the order $N(r,n,m,k)$ of the variety Φ_r, we determine the number of points of free

[1] For the derivation of this formula see Bertini, *Projektive Geometrie Mehrdimensionaler Räume*, 1924, Kapitel XII.

[2] "A series of rational loci with one apparent double point," *Proceedings of the Cambridge Philosophical Society*, vol. 27 (1931), pp. 399-403.

[3] "On a certain rational V_n^{2n+1} in S_{2n+1}," *American Journal of Mathematics*, vol. 56 (1934), pp. 219-224.

[4] "The number of apparent double points of certain loci," *Proceedings of the Cambridge Philosophical Society*, vol. 28 (1932), pp. 285-299.

intersection of r variable hypersurfaces of $|V|$. We may consider, without loss of generality for our purpose, r composite hypersurfaces of $|V|$ each composed of a hypersurface U of order m containing $M^{m^k}_{r-k}$ and a general hypersurface W of order $n-m$ which is in no way related to $M^{m^k}_{r-k}$. The $N(r,n,m,k)$ points of intersection of these r composite V's can now be calculated without difficulty. First, we have the $(n-m)^r$ points common to all the r W's. Then, we see that the U belonging to any one of the r V's and the W's belonging to the remaining $r-1$ V's intersect in $m(n-m)^{r-1}$ points. There are in all $\binom{r}{1}m(n-m)^{r-1}$ such points. Again, the U's of any two of the r V's and the W's of the other $r-2$ V's intersect in $m^2(n-m)^{r-2}$ points. The total number of such points is $\binom{r}{2}m^2(n-m)^{r-2}$. If we continue in this manner, we shall soon arrive at the final step where we find that the U's of any $k-1$ of the r V's and the W's of the remaining $r-k+1$ V's meet in $m^{k-1}(n-m)^{r-k+1}$ points. We have in all $\binom{r}{k-1}m^{k-1}(n-m)^{r-k+1}$ such points. Since there are no other points common to the r composite V's outside of the basic variety, we have, adding,

$$N(r,n,m,k)=\sum_{i=0}^{k-1}\binom{r}{i}m^i(n-m)^{r-i} \tag{2}$$
$$=n^r-\sum_{i=0}^{r-k}\binom{r}{i}m^{r-i}(n-m)^i$$

for the order of the variety Φ_r.

In a similar manner, we find that the order of the $(r-q)$-dimensional variety V_{r-q} of intersection of q variable hypersurfaces of $|V|$ is $N(q,n,m,k)$ whose value is obtainable from (2) by replacing r by q. If q is increased by unity, we have a V_{r-q-1} of intersection of $q+1$ hypersurfaces of $|V|$, of order $N(q+1,n,m,k)$. We now find the intersection of V_{r-q} and the base manifold $M^{m^k}_{r-k}$. For this purpose, we notice that V_{r-q} and another hypersurface of $|V|$ intersect completely in a composite $(r-q-1)$-dimensional variety of order $n\cdot N(q,n,m,k)$. One of the components is the V_{r-q-1} just mentioned and the other, of order $n\cdot N(q,n,m,k)-N(q+1,n,m,k)$, lies on $M^{m^k}_{r-k}$ and is therefore the intersection of V_{r-q} and $M^{m^k}_{r-k}$.

3. *Varieties corresponding to sub-spaces of S_r.* A general $(k-t)$-space, where $1\leqq t\leqq k$, in S_r has no point in common with $M^{n^k}_{r-k}$ and therefore meets $|V|$ in a system of $(k-t-1)$-dimensional varieties of order n having no base manifold. The dimenison of this system is, then,

$$\rho(k-t,n,m,k)=\binom{n+k-t}{k-t}-1.$$

Therefore, the transform of S_{k-t} is a $\Phi^{n^{k-t}}_{k-t}$ of order n^{k-t} on Φ_r, contained in a space of $\binom{n+k-t}{k-t}-1$ dimensions. Thus, for $t=k$, a point in S_r is the image

of a point on Φ_r; for $t = k-1$, a line of S_r is the image of a curve of order n contained in an n-space; and so on.

If we consider a general $(k+h)$-space S_{k+h}, where $0 \leqq h \leqq r-k$, in S_r, we see that it meets M^{mk}_{r-k} in a manifold $M_h{}^{mk}$ and meets $|V|$ in a $\rho(k+h, n, m, k)$-dimensional system of $(k+h-1)$-dimensional varieties of order n having $M_h{}^{mk}$ for base manifold. The transform of S_{k+h} is a Φ_{k+h} of order $N(k+h, n, m, k)$ on Φ_r, contained in a $\rho(k+h, n, m, k)$-space. Thus, for $h=0$, an S_k has for transform a Φ_k of order $N(k, n, m, k) = n^k - m^k$; for $h=1$, an S_{k+1} has for transform a Φ_{k+1} of order

$$N(k+1, n, m, k) = n^{k+1} - m^{k+1} - (k+1)m^k(n-m); \text{ etc.}$$

From the fact that a general S_k meets M^{mk}_{r-k} in points we see that a point of M^{mk}_{r-k} is the image of a $(k-1)$-space on Φ_r. Then, the $M_h{}^{mk}$ common to S_{k+h} and M^{mk}_{r-k} is the image of a $(k+h-1)$-dimensional manifold μ_{k+h-1} of order $\nu(k+h, nm, k)$, a locus of ∞^h $(k-1)$-spaces. We now calculate $\nu(k+h, n, m, k)$. Consider an S_{k+h-1} of S_{k+h}. Its transform on Φ_r is a Φ_{k+h-1} of order $N(k+h-1, n, m, k)$. Therefore, a $V^n{}_{k+h-1}$ of order n in S_{k+h} will have for transform a variety of order $n \cdot N(k+h-1, n, m, k)$. Now let $V^n{}_{k+h-1}$ pass through $M_h{}^{mk}$, and the transform is now of order

$$n \cdot N(k+h-1, n, m, k) - \nu(k+h, n, m, k).$$

But this order is equal to $N(k+h, n, m, k)$ which is that of the section of Φ_{k+h} by a $\Sigma_{\rho(k+h,n,m,k)-1}$ of $\Sigma_{\rho(k+h,n,m,k)}$. Then, we have

$$\begin{aligned}\nu(k+h, n, m, k) &= n \cdot N(k+h-1, n, m, k) - N(k+h, n, m, k)\\ &= \binom{k+h-1}{h} m^k (n-m)^h.\end{aligned}$$

Thus, for $h = r-k$, the transform of M^{mk}_{r-k} is a μ_{r-1} of order

$$\nu(r, n, m, k) = \binom{r-1}{r-k} m^k (n-m)^{r-k}.$$

If $h=1$, we see that the curve $M_1{}^{mk}$ in which an S_{k+1} meets M^{mk}_{r-k} is the image of a μ_k of order $km^k(n-m)$; and, if $h=0$, each of the m^k points common to an S_k and M^{mk}_{r-k} is the image of a $(k-1)$-space.

4. *Rational systems of loci on* Φ_r. A hyperplane, S_{r-1}, of S_r, as was seen above, goes into a variety of order $N(r-1, n, m, k)$, and hence an n'-ic hypersurface, where $m \leqq n' \leqq n$, goes into one of order $n' \cdot N(r-1, n, m, k)$. Consider a $V^{n'}_{r-1}$ passing through M^{mk}_{r-k}, and there are $\infty^{\rho(r,n',m,k)}$ such hypersurfaces. Now the transform of this $V^{n'}_{r-1}$ is a $\Phi^{N'}_{r-1}$ of order

$$N' = n' \cdot N(r-1, n, m, k) - \binom{r-1}{r-k} m^k (n-m)^{r-k}.$$

For the various values of n' from m to n, there are, then, various rational $\infty^{\rho(r,n',m,k)}$-systems of $(r-1)$-dimensional varieties of order N'. If, in particular, we put $n'=m$, $k=2$, we see that Φ_r contains a rational pencil of varieties of dimension $r-1$ and order

$$N'=m\cdot N(r-1,n,m,k)-(r-1)m^2(n-m)^{r-2}=m(n-m)^{r-1}.$$

5. *The case* $k=2$, $n=m+1\geqq 3$. In this interesting case we have a system $|V|$, of dimension[5] $\rho=2r+1$, of hypersurfaces of order n passing through the complete intersection $M_{r-2}^{(n-1)^2}$ of two general hypersurfaces of order $n-1$, in S_r. The equation of a general member of the system is of the form

$$(3)\qquad F_0\cdot\sum_{i=0}^{r}A_ix_i+F_1\cdot\sum_{i=0}^{r}B_ix_i=0,$$

where $F_0=0$, $F_1=0$ are the equations, of degree $n-1$ in $x_0,x_1,\cdots,x_r$, representing the two hypersurfaces intersecting in $M_{r-2}^{(n-1)^2}$ and the A's and B's are arbitrary constants. Setting

$$(4)\qquad X_{00}:X_{01}:\cdots:X_{0r}:X_{10}:\cdots:X_{1r}$$
$$=F_0x_0:F_0x_1:\cdots:F_0x_r:F_1x_0:\cdots:F_1x_r,$$

where X_{ij} $[i=0,1;\ j=0,1,\cdots,r]$ are the coördinates of a point in a $(2r+1)$-space Σ_{2r+1}, we have the equations of the variety Φ_r^N, where $N=rn-r+1$, represented upon S_r by the hypersurfaces of $|V|$.

A hyperplane in S_r, say $x_r=0$, is transformed by (4) into a $\Phi_{r-1}^{N'}$ of order $N'=rn-n-r+2$, which is of the same nature as Φ_r^N for r diminished by unity. Then, the transform of a hypersurface of order $n-1$ is a variety of order $N'(n-1)$. If the hypersurface is one, say V_{r-1}^{n-1}, of the pencil determined by $F_0=0$, $F_1=0$, the corresponding variety degenerates into the transform μ_{r-1} of $M_{r-2}^{(n-1)^2}$, of order $(r-1)(n-1)^2$, and a Θ_{r-1}^{n-1} of order $n-1$, which is the proper transform of V_{r-1}^{n-1}. Thus, Φ_r^N contains a rational pencil of $(r-1)$-dimensional varieties of order $n-1$ corresponding to the pencil $F_0+\lambda F_1=0$ in S_r.

6. *Different varieties on* Φ_r^N. Each of the ∞^1 Θ_{r-1}^{n-1}'s on Φ_r^N is contained in an r-space and the ∞^1 containing r-spaces form a rational locus, Ω_{r+1}^{r+1}, of order $r+1$ whose equations are

[5] For $k=2$, $n=2$, $m=1$, the system of quadric hypersurfaces in r-space having a fixed $(r-2)$-space in common is of dimension $2r$ and not $2r+1$. This system can best be regarded as a special case of the linear system of hypersurfaces of order n passing through a given $(r-2)$-space $n-1$ times.

D. W. Babbage, in the paper cited in footnote [2], has already studied the case $n=3$.

$$\begin{Vmatrix} X_{00} & X_{01} & \cdots & X_{0r} \\ X_{10} & X_{11} & \cdots & X_{1r} \end{Vmatrix} = 0.$$

A general $(r+2)$-space of Σ_{2r+1} meets Φ_r^N in a curve Γ^N and Ω^{r+1}_{r+1} in a ruled surface of order $r+1$ whose rulings are all $(n-1)$-secant lines of the curve.

Now Φ_r^N contains $\infty^{(t+1)(r-t)}$ Φ_t's of order $tn-t+1$ each contained in a $(2t+1)$-space. They correspond to the $\infty^{(t+1)(r-t)}$ t-spaces of S_r and are of the same nature as Φ_r^N for $r=t$. The $(2t+1)$-space of a Φ_t meets Ω^{r+1}_{r+1} in an Ω^{t+1}_{t+1}. This Ω^{t+1}_{t+1} is the locus of ∞^1 t-spaces and each of these t-spaces meets a Θ^{n-1}_{r-1} in a Θ^{n-1}_{t-1}. Therefore, Φ_t contains ∞^1 Θ^{n-1}_{t-1}'s.

Consider an $(r+t+1)$-space Σ_{r+t+1} passing through the $(2t+1)$-space Σ_{2t+1} of a Φ_t. It meets Ω^{r+1}_{r+1} in a $(t+1)$-dimensional variety of order $r+1$. Since Σ_{2t+1} already meets Ω^{r+1}_{r+1} in an Ω^{t+1}_{t+1}, the variety just mentioned is composed of Ω^{t+1}_{t+1} and $r-t$ $(t+1)$-spaces. Note that each of these $(t+1)$-spaces contains a Θ_t^{n-1}.

Now a general $(r+t+1)$-space meets Φ_r^N in a t-dimensional variety of the same order. If the $(r+t+1)$-space passes through Σ_{2t+1}, it contains Φ_t which is of order $tn-t+1$. Then, its intersection with Φ_r^N is composed of Φ_t and $r-t$ Θ_t^{n-1}'s.

We may make use of the preceding result to find the genus of the curve Γ^N in which an $(r+2)$-space meets Φ_r^N. Putting $t=1$, we have a composite curve whose components are a Φ_1^n which is a rational curve of order n and $r-1$ Θ_1^{n-1}'s which are plane curves of order $n-1$. Each Θ_1^{n-1} has $n-1$ points in common with Φ_1^n and therefore has $(n-1)^2$ apparent intersections with it. There are $(r-1)(n-1)^2$ apparent intersections between Φ_1^n and the $r-1$ Θ_1^{n-1}'s; there are also $\frac{1}{2}(r-1)(r-2)(n-1)^2$ apparent intersections of the $r-1$ Θ_1^{n-1}'s two by two. Φ_1^n itself has $\frac{1}{2}(n-1)(n-2)$ apparent double points. Therefore, we say that the total curve Γ^N has

$$\begin{aligned}(r-1)(n-1)^2+\tfrac{1}{2}(r-1)(r-2)(n-1)^2+\tfrac{1}{2}(n-1)(n-2)\\ =\tfrac{1}{2}(n-1)[(r^2-r+1)n-(r^2-r+2)]\end{aligned}$$

apparent double points. Denote this number by b_{r-1}. We note that b_{r-1} is also the order of the double variety, $\Delta^{b_{r-1}}_{r-1}$, on the projection of Φ_r^N upon an $(r+1)$-space. The genus of Γ^N is, then,

$$\tfrac{1}{2}(N-1)(N-2)-b_{r-1}=\tfrac{1}{2}(r-1)(n-1)(n-2).$$

On Φ_r^N we have ∞^{2r-2} curves of order n and they are rational curves and correspond to the ∞^{2r-2} lines of S_r. Each of them is contained in a 3-space and this 3-space meets Ω^{r+1}_{r+1} in a quadratic regulus Ω_2^2. Of the ∞^{2r-2} such

3-spaces one and only one passes through a general given point P of the $(2r+1)$-space Σ_{2r+1} containing Φ_r^N. Therefore, through P we can construct $\frac{1}{2}(n-1)(n-2)$ lines meeting in two points the rational curve of order n of Φ_r^N contained in the 3-space through it, and these are the only lines through P bisecant to Φ_r^N. Hence, we say that Φ_r^N has $\frac{1}{2}(n-1)(n-2)$ apparent double points and that its projection upon a $(2r)$-space of Σ_{2r+1} has $\frac{1}{2}(n-1)(n-2)$ improper double points.

7. *Projections and sections of* Φ_r^N. Let us project Φ_r^N from an $(r-1)$-space, denoted by Z_{r-1}, of Σ_{2r+1} upon an $(r+1)$-space Σ_{r+1}. Denote the projection by ${}_{r-1}\Phi_r^N$. We have already remarked, in § 6, that the double variety $\Delta_{r-1}^{b_{r-1}}$ on ${}_{r-1}\Phi_r^N$ is of order $b_{r-1}=\frac{1}{2}(n-1)[(r^2-r+1)n-(r^2-r+2)]$. Now we verify this fact. Suppose the center of projection, Z_{r-1}, is contained in the Σ_{2r-1} containing a $\Phi_{r-1}^{rn-n-r+2}$ [see § 6]. This Σ_{2r-1} meets Ω_{r+1}^{r+1} in an Ω_r^r which is in turn met by Z_{r-1} in r points. Designate these points by $A^{(1)}, A^{(2)}, \cdots, A^{(r)}$. Through $A^{(i)}$ passes an r-space $\Sigma_r^{(i)}$ containing a Θ_{r-1}^{n-1}. Now the projection ${}_{r-1}\Phi_r^N$ in Σ_{r+1} contains an $(rn-n-r+2)$-fold $(r-1)$-space which is the projection of $\Phi_{r-1}^{rn-n-r+2}$ in Σ_{2r-1} and r $(n-1)$-fold $(r-1)$-spaces which are the projections of the Θ_{r-1}^{n-1}'s in the Σ_r's through $A^{(1)}, \cdots, A^{(r)}$. The $(rn-n-r+2)$-fold $(r-1)$-space is equivalent to a double variety of order $\frac{1}{2}(rn-n-r+r)(rn-n-r+1)$ and each of the r $(n-1)$-fold $(r-1)$-spaces is equivalent to a double variety of order $\frac{1}{2}(n-1)(n-2)$. Therefore, the total double variety is of order

$$b_{r-1}=\tfrac{1}{2}(rn-n-r+2)(rn-n-r+1)+\tfrac{1}{2}r(n-1)(n-2)$$

which is reduced to the value already given.

We now find the pinch variety, $J_{r-2}^{c_{r-2}}$, on the projection ${}_{r-1}\Phi_r^N$. We remark that the $\Phi_{r-1}^{rn-n-r+2}$ in the Σ_{2r-1} containing Z_{r-1} has an apparent double variety of order $b_{r-2}=\frac{1}{2}(n-1)[(r^2-3r+3)n-(r^2-3r+4)]$ and is therefore of rank

$$a=(rn-n-r+2)(rn-n-r+1)-2b_{r-2}=(n-1)(rn-2n+2).$$

Then, on the $(rn-n-r+2)$-fold $(r-1)$-space, the projection of $\Phi_r^{rn-n-r+2}$, is a pinch variety of order a. Now each of the r Θ_{r-1}^{n-1}'s is of rank $(n-1)(n-2)$ and, therefore, in each of the r $(n-1)$-fold $(r-1)$-spaces, the projections of the Θ_{r-1}^{n-1}'s, is a pinch variety of order $(n-1)(n-2)$. Then, the total locus of pinch points is of order $c_{r-2}=a+r(n-1)(n-2)=2(r-1)(n-1)^2$.

Consider a general projection of Φ_r^N upon a Σ_{r+1}, that is, one from a general center of projection upon a general $(r+1)$-space. Denote it also by ${}_{r-1}\Phi_r^N$. There are ∞^1 r-spaces in Σ_{r+1} each meeting ${}_{r-1}\Phi_r^N$ in a composite

section composed of a variety of order $rn - n - r + 2$ and one of order $n - 1$. These ∞^1 r-spaces envelop a curve of order $r + 1$. Let a 3-space meet ${}_{r-1}\Phi_r^N$ in a surface F^N. The ∞^1 planes containing composite sections envelop a curve of order $3(r-1)$, rank $2r$, class $r + 1$, with $4(r-2)$ cusps.

8. *Other projections of* Φ_r^N. Suppose we now project Φ_r^N from an $(r-2)$-space, Z_{r-2}, upon an $(r+2)$-space Σ_{r+2}. The projection, ${}_{r-2}\Phi_r^N$, has a locus, $\Delta_{r-2}^{b_{r-2}}$, of double points, of order b_{r-2}, and a locus, $J_{r-3}^{c_{r-3}}$, of pinch points, of order c_{r-3}. A 4-space section of the projection is a surface with b_{r-2} improper nodes. By making use of a relation, which is known concerning surfaces in 4-space,[6] namely, $2b_{r-1} - 2b_{r-2} = c_{r-2}$, we find that

$$\begin{aligned} b_{r-2} &= b_{r-1} - \tfrac{1}{2}c_{r-2} \\ &= \tfrac{1}{2}(n-1)[(r^2 - r + 1)n - (r^2 - r + 2)] - (r-1)(n-1)^2 \\ &= \tfrac{1}{2}(n-1)[(r^2 - 3r + 3)n - (r^2 - 3r + 4)]. \end{aligned}$$

This value is identical with that found in § 7 for the double variety $\Delta_{r-2}^{b_{r-2}}$ on the projection of $\Phi_{r-1}^{rn-n-r+2}$ upon an r-space. To verify this fact, put Z_{r-2} in the Σ_{2r-1} containing $\Phi_{r-1}^{rn-n-r+2}$. Then, the $\Delta_{r-2}^{b_{r-2}}$ on the projection ${}_{r-2}\Phi_r^N$ is the same as that on the projection of $\Phi_{r-1}^{rn-n-r+2}$. From this last fact we derive the result that the order of $J_{r-3}^{c_{r-3}}$ is $c_{r-3} = 2(r-2)(n-1)^2$.

Reasoning in the same manner, we see that the projection ${}_{r-t}\Phi_r^N$ from a Z_{r-t} of Σ_{2r+1} upon an $(r+t)$-space has a double locus $\Delta_{r-t}^{b_{r-t}}$ of order $b_{r-t} = \tfrac{1}{2}(n-1)[(r^2 - 2rt + t^2 + r - t + 1)(n-1) - 1]$ and a pinch locus $J_{r-t-1}^{c_{r-t-1}}$ of order $c_{r-t-1} = 2(r-t)(n-1)^2$. For $t = r$, the projection ${}_0\Phi_r^N$ upon Σ_{2r} has $b_0 = \tfrac{1}{2}(n-1)(n-2)$ improper double points, as was found in § 6.

9. *Various special cases.* It is of interest to mention that, if $r = n$, we have a $\Phi_n^{n^2-n+1}$ in a Σ_{2n+1}. It is ruled, being the locus of ∞^{n-1} lines, which correspond to the $(n-1)$-secant lines of $M_{n-2}^{(n-1)^2}$ in S_n. The number of lines passing through each point of the variety is $(n-1)!$ What has just been said holds true for $n = 2$, except that the ruled cubic surface Φ_2^3 is in a 4-space and not in a 5-space. If $r \geqq n$, then Φ_r^N contains ∞^{2r-n-1} lines of which ∞^{r-n} pass through each point. If $r \leqq n$, Φ_r^N may be regarded as a locus of ∞^{r-1} curves of order $n - r + 1$. Through a general point pass $(r-1)!\binom{n-1}{r-1}^2$ of them.

THE UNIVERSITY OF CALIFORNIA.

[6] Severi, "Intorno ai punti doppi impropri di una superficie generale dello spazio a quattro dimensioni, e a'suoi punti tripli apparenti," *Rendiconti di Palermo*, vol. 15 (1901), pp. 33-51.